

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 27, 2001, 16:54:05 ; Search time 7648.39 Seconds

(Without alignments)
5996.315 Million cell updates/sec

Title: US-09-830-647-3

Perfect score: 2780

Sequence: 1 aatttcgacgcagcgtctctg.....aaaaaaaaaaaactcgaag 2780

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vi: *
30: em_htgo_hum: *
31: em_htgo_in: *
32: em_htgo_rod: *
33: em_htg_hum: *
34: em_htg_in: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2780	100.0	2780	9	AB028069	AB028069 Homo sapi
2	2648	95.3	2719	9	AB028070	AB028070 Homo sapi
3	2458.8	88.4	2474	9	AF160249	AF160249 Homo sapi
4	2456	88.3	2460	9	AF160876	AF160876 Homo sapi
5	2118.6	76.2	236236	2	AC018977	AC018977 Homo sapi
6	1537	55.3	2458	10	AF292400	AF292400 Cricetulu
7	1236	44.5	2309	10	MMDBF4	MMDBF4
8	1194.2	43.0	123331	9	AC005164	AC005164 Mus muscu
9	1194.2	43.0	163031	2	AC024947	AC024947 Homo sapi
10	1194.2	43.0	163889	2	AC004958	AC004958 Homo sapi
11	1101.6	39.6	2276	9	AK022969	AK022969 Homo sapi
12	1003.8	36.1	232903	2	AC091341	AC091341 Rattus no
13	554.6	19.9	136823	2	AC003083	AC003083 Homo sapi
14	513.8	18.5	685	6	AX070167	AX070167 Sequence
15	423.6	15.2	177899	2	AC074175	AC074175 Mus muscu
16	414.4	14.9	433	11	G36582	G36582 SHGC-53468
17	411	14.8	422	6	AX072042	AX072042 Sequence
18	384.4	13.8	403	6	AX070196	AX070196 Sequence
19	332.6	12.0	205662	9	AC017002	AC017002 Homo sapi
20	327	11.8	136555	2	AC068279	AC068279 Homo sapi
21	327	11.8	157063	2	AC083899	AC083899 Homo sapi
22	217.4	7.8	220	11	G41942	G41942 SHGC-64833
23	211.4	7.6	397	6	AX070116	AX070116 Sequence
24	126.2	4.5	168717	2	AC024683	AC024683 Homo sapi
25	126.2	4.5	177720	9	AC007514	AC007514 Homo sapi
26	104.6	3.8	1517	9	AF125531	AF125531 Homo sapi
27	99.4	3.6	377	6	AX070662	AX070662 Sequence
28	67.8	2.4	1141	6	AX083744	AX083744 Sequence
29	65.6	2.4	1141	6	AX083744	AX083744 Sequence
30	64.2	2.3	87256	2	AC013754	AC013754 Homo sapi
31	63.6	2.3	207420	2	AC078884	AC078884 Mus muscu
32	63.4	2.3	321003	2	PFMALP3	PFMALP3
33	62	2.2	7218	6	I64494	I64494 Sequence 14
34	61.2	2.2	71559	2	AC091013	AC091013 Homo sapi
35	61	2.2	245802	2	AC006279	AC006279 Plasmodu
36	60.8	2.2	51274	2	AC023454	AC023454 Homo sapi
37	60.8	2.2	167968	2	AC084398	AC084398 Homo sapi
38	60.6	2.2	140176	2	AC016411	AC016411 Homo sapi
39	60.6	2.2	174133	8	AF083031	AF083031 Homo sapi
40	60.2	2.2	201676	2	AC090647	AC090647 Mus muscu
41	58.8	2.2	120091	2	AC012684	AC012684 Homo sapi
42	59.8	2.2	170102	9	AC008079	AC008079 Homo sapi
43	59.6	2.1	101509	2	AC027353	AC027353 Homo sapi
44	59.4	2.1	62164	2	AC068512	AC068512 Homo sapi
45	59.4	2.1	158167	2	AC034105	AC034105 Homo sapi

ALIGNMENTS

RESULT	1	AB028069	2780 bp	mrna	PRI	03-JUN-1999
LOCUS	AB028069	Homo sapiens mRNA for activator of S phase kinase, complete cds.				
DEFINITION	AB028069	Homo sapiens mRNA for activator of S phase kinase, complete cds.				
ACCESSION	AB028069.1	GI:4996095				
VERSION	AB028069.1	GI:4996095				
KEYWORDS		activator of S phase kinase.				
SOURCE		Homo sapiens Hela cDNA to mRNA, clone:H37.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS		1 (sites)				
TITLE		Araki, K. and Masai, H., Yamada, M., Mahony, D., Seghezzi, W., Lees, E.,				
JOURNAL		A Novel Growth- and Cell Cycle-Regulated Protein, ASK, Activates Human Cdc7-Related Kinase and is Essential for G1/S transition in Mammalian Cells				
REFERENCE		Mol. Cell. Biol. (1999) In press				
AUTHORS		2 (bases 1 to 2780)				
		Masai, H. and Kumagai, H.				

TITLE Direct Submission
JOURNAL Submitted (28-MAY-1999) to the DBJ/EMBL/GenBank databases. Hisao Mesai, Institute of Medical Science, University of Tokyo, Department of Developmental Biology, 4-6-1 Shirokane-dai, Minato-ku, Tokyo 108-8639, Japan (E-mail: hisao@ems.u-tokyo.ac.jp, Tel: 01-3-5449-5661, Fax: 01-3-5449-5424).

FEATURES	Location/Qualifiers
source	1. .2780

```

gene      518.  .2542
          /gene="ASK/H37 (hubb4)"
CDS       518.  .2542

```

"/function="binds and activates hucbcr7 kinase"/
/note="ASK encodes a regulatory subunit for hucbcr7, the human homologue of budding yeast GclC kinase." Amino acid 51-90 and 280-322 are motif-N and motif-C, respectively conserved in bDifA-related molecules. Amino acid 201-216 is a putative bipartite nuclear localization signal. Amino acid 101-120 and 552-564 are two putative PSTT-like sequences."

```

/codon_start=1
/product="activator of S phase Kinase"
/protein_id="BA078326.1"
/db_xref="GI:4960965"
/translation="MNSGAKRHSGHPGGIOVKNENRPSLKSLTQDNBPBCKSC
PLNGKVEYLDPSVYISKRLQKDIQDGGREVEEFLSTVLSNNKKEKPKQGTGR
SPVSPESAVYAEFTTSPHSDGSPKSDTSLRGSLVYERKIDHPTISNLSL
NALSMGSLIKLIDHDIARYIDOKKREYLKLSKTSYVRGCKRAGCAQTRGRLK
HYVEVMGSLVYPIYLTQNMPTINYSQIQSPQDNPBEMKQVYKILQYDGGK
KYGCSLIDQLKKKKKGYCECLQYELPHLLSEPHNRQMSQNYVDYIDSKY
VPPFVYERKDPKKKKRIKYSGSLSPVYSAYLKTDEQKVEYLDHISQCCQEDTIT
KEQNLFTKEQTEKRLFTSEFPHPNSELNGLKQNSVNSCMLSTAEEDLRQNPY
LPHAKKQCLIDLSERTLSNDLEELRDHDKTQIAKSNVSDVSDTNGSQQPQKQ
DYLPAFAKQKDEKSHLFTHDSGLTITNSDEHLYVAKPFTPEENEDENRQK
DLSAPSKTHRWKYLILGNKRNELPNAEFDRTSEFTLTQENRILCSSPVASLDFQY
SEKSESELEFSLYERKSGICNVNLDLWEENSDNLLTAFFSSPSTFTG?"
2724..2730
polyA_signal

```

Query Match	100.0%	Score 2780	DB 9	Length 2780
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2780	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	aattggcaagagctctctctgagctgcgcgcaagcctgaagcggcggagccgagcccg	60	
Db	1	AATTGGCAACAGCTCTCTGAGCTGCGCAAGCTTGAAGCGGACCGAGACCCCG	60	
QY	61	gtctgagactgagagagcaacggaatbgaaggcggggttagagcgaggaacaacactcgag	120	
Db	61	GTCTGAGACTAGACGACCAACGGAATGGAGCGGGGTGAGCGCGCAAAACCAACTCTGAG	120	
QY	121	ggccagagacgagcgcgcgagaaagagacggcgcgctgaaggggcgagggcgcgcaagcgagaa	180	
Db	121	GGCCAGAGACGAGCGCGCGAGAGACGCGCGGTGAGGCGCGGCGCGCCACGCGGGA	180	
QY	181	ggccaagacagagggggcggaagcggaagcgagggcgagggcgctgagctgtaaaacggggcgagg	240	
Db	181	GGCAGACGACGAGGGGCGCAGCGCAGGCGGCGCGCTGGCTGTAACGCGGCGGGG	240	
QY	241	ggcgcgatctggcgcgcgcgcgcgctgaagcgcttcaacttcaacccgcgcgaagccca	300	
Db	241	GGCGCGATCTGGCGCGCGCGCGCGCTGAAGCGCTTTCAAACTTCAACGCGCGACGCCA	300	
QY	301	ctcgcttggctgttgcgacttccctccctcgacttgaagcggatccggcccgcccggaac	360	

Db	301	CTGCTTTTGCTTTTCGGCCCTTCCTCCGCGCGCTGTGGAGCGGATCGGCGCCCGGCAAC	360
Qy	361	CGAAGCTGAGAGCGGTAAGTCTCTAATCGTGTAGAGCGGTAGCTGCGAGAGAGAGAG	420
Db	361	CGAGCTGAGAGCGCGGTACTCTACTGCGGTAGAGCGGCTGACTGCGGAGAGAGAG	420
Qy	421	CGGCGTCTGTCAACAGCGGAGGAGAAAGCGTGTCTTCGGAGTGGCGGTGCGACAT	480
Db	421	CGGCGCTCTGTCAACAGCGCGGAGAAAGCGCTTTCGGCGGTGCGCGTCTGACAT	480
Qy	481	ttctccggagcccaagatgtagtggccgggagatcgatgaacccggagacatgagat	540
Db	481	TTCTCCGGAGCCCAAGATGTAGGTATCGGGGAGTGTGATGATCCGAGACATGAGAT	540
Qy	541	ccacagtaaaggacatctccagggtggaaatccaaagtcacaaatgaaaaaacagacatc	600
Db	541	CCACAGTAAAGGACATCTTCCAGGGTGGATTCACAGTCACAAATATATAAAAAACAGACATC	600
Qy	601	tctgaaatctctgaaacatgataacagggccggaaataatccaaatgtaagccacttggg	660
Db	601	TCGAAATCTCTGAAATGATATTAACAGCCGCAAAAATTCAAATGTAAAGCCATTTGGG	660
Qy	661	aaagatattcttccttgacttaacctctgtccacatctcgtaaaacatccaaaggacat	720
Db	661	AAAGATATTTTACCTTGACTTACCTCTGTGCACATCTGTAAAACTTAAAAAGACAT	720
Qy	721	taagagctcggagagggaggttgaaagatcttcagcaaaagataatcagttatctcttc	780
Db	721	TAAGATCTGGAGGGCGAGTTGAAATTTCTCAGCAAAATATTCAGTTATCTTATTTTC	780
Qy	781	aaataaagagagagtaaatatgcaaaaaccttggatcgaaattcttcctgtaccaagctc	840
Db	781	AATATAGAGGAGAGCTAAATTTGGACAAACCTTGGGTGGATTTCTCTGTACCAAGTCC	840
Qy	841	agaaatgataatctctgcaaaaacaccttcaacctccagatccggacatatgaaatcat	900
Db	841	AGAAATGTGATTAAGTGTGCAAAACACCTTCACTCATCCAGCCATATGGAAGTTCAAT	900
Qy	901	taagttacagacacaaagtgtgtttaagcagggaaaatatatgataaaagctctcaa	960
Db	901	TAAGTATACAGACAAAGGTGTGTTTAAACAGAGAAAATTTATGTTTAAAAACCTATCAA	960
Qy	961	ggagcagatgattatctccatccaaatgatatataccaatgacctgtgtcatgvgggataa	1020
Db	961	GGAGCAGTGTATTTATCTCTTCAAAATGTATTTATTCAAATGCTGTGCAATGGGAGATAA	1020
Qy	1021	aattcttcatactgcatgacatagatacaatcgacaacaaagaaaagatgtaattc	1080
Db	1021	AATTTCTCATTTGATGATGACATTAAGATATCATTAATACAAAAAGAAAAGATGTGATTT	1080
Qy	1081	actcaagaatatcaatctatcttcagtaagagaatgggggcaaaagatgtgtgtgtgcaca	1140
Db	1081	ACTCAAGAAATCAAGTACTTCTACAGTAAGACATGGGGCAAAAAGACTTGTGTGTGTGCACA	1140
Qy	1141	aaaaaacaagaacaggaagactcacaagaagcctttgtaaaggtggaagatactgagcaact	1200
Db	1141	AAAAAACAAACAGGAACAGCAACCAAAAAGCCTTTGTAAAGGTGGAAATATGAGCCAACT	1200
Qy	1201	ttaagagcaattatcttcagctgcagccaaatgtgctttttaaattatctctctcaaga	1260
Db	1201	TTTATAGCCATTTTATCTTCAGCGACCAATATGCGCTTTTAAATATTTCAATTGACAGA	1260
Qy	1261	ggccctcagctcaattgattagacaagcactctgtatgaaagagaaactcaagttaa	1320
Db	1261	GCCCTCAAGTCAATTGATGTATGTATGACAGCCACTGTGTGCAAAAGCAAACTCAGTTAA	1320
Qy	1321	actaagaatccaaacagatggcgataagatagtgvgaaacctcaatcaactcagttgaa	1380
Db	1321	ACTAAATATTCACAAAGATGTGCGCATTAAGATGTGTGGAACCTCAATTCAGTTGAA	1380
Qy	1381	aggaagaagaaaagagatattgtgtaattgttcgtcttgacgaataatgaaatctagaac	1440
Db	1381	AGGAAGAAGAAAAGAGATATTGTGTAATGTTGCTTCAGCAAAATATTCAGATCTAGAAAC	1440

OY 1441 tcaactctaaatgagcaacacagaacactttgcacagagtaacagatcatcagtttgtga 1500
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1441 TCACCTTTAACTGAGCAGACAGAAACTTTGACAGAGTAAACCAATCAAGTTGTTGA 1500
 OY 1501 tgaatactatcctaagttagtttctgactttgttggaataagaagaacacacttaaaa 1560
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1501 TGATATTTGATCTAAGTTAGTTTGTGACTTTGTGGAATGTGAAGAGACACACTAAAAA 1560
 OY 1561 gaaaagaataaataatcagttgttgatcccttctccctgttctgcgaagtgtccctga 1620
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1561 GAAAAAATTAATACAGAGTTGATCCCTTCTCTGTTTGTGCAAGTCTCTGAAAAA 1620
 OY 1621 gactgaacaaaaggaagatggaatttgcacatatcttcgaaagatttgcagaaga 1680
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1621 GACTGACAAAGGAAAAAGTGGAAATTCACAAATTTCTCAGAAAGATTGCCAGAGAGA 1680
 OY 1681 tgatacaacagtgagagagcagaatttccctgtataaagagaccggaacatgaaaaa 1740
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1681 TGATACACAGTGAAGAGAGCAGAAATTTCTGTATGAAGACCCAGGAACCTGAAAAA 1740
 OY 1741 gctcctgtttattcagaagccatcccccaccccttcaaatgaatgagaggtctaa 1800
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1741 GCTCCTGTTTATTTAGAGCCCATCCCCACCTTCAAAATGAATGAGAGGCTTAATGA 1800
 OY 1801 gaaaatgagtaataaatttccatgtttaagtaagctgagatgagatgaagaagattt 1860
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1801 GAAATATGATTAATTAATGTTCCATGTTAACTGTAAGTACAGTGAAGATGACATGAAGATTT 1860
 OY 1861 tacaacagctacccttcatataaacaacaggaatgacatttctgacatttccgaacac 1920
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1861 TACACAGCTACTCTACATATAAATAACAAACAGATGATCTTGATGACTTTCCGAAACAC 1920
 OY 1921 attaaatgaaaaatgacttagaagaactaagggtagatcatcataatgaatgaacagc 1980
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1921 ATTAAGTGAATGACTTGAAGAACTAAGGTAGATCACTAATTAATGATACATACAGC 1980
 OY 1981 atctgtacatgttctcgtatcattcaatgacagataatagtgatcttcaacccaagaagatc 2040
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1981 ACTGTACATGTTTCTGATTCAGATTGATACATATATGATGATCTCAACCAACAGAAAGTC 2040
 OY 2041 aataactgacttcttccagcaagaagatcagaagaagaagccttcatcatatttcc 2100
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2041 AATATCTGCTGCTTTTCCAGCAAGATCTCAAGCAAAAGACCTTCATTCATATTTTAC 2100
 OY 2101 tcaatgattcgtcttgatacaataaacaagttcacagaagacacctaactgttcaagcaca 2160
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2101 TCATGATTTCTGCTGATTAACAATTAACAGTTTCAAGAGCACACTTAAGTGTTCAGGCATA 2160
 OY 2161 ggtctcatcccaactcctccctcgagagaaacccaatgaatgtgacttaagaataatgata 2220
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2161 GGCTCCATTCCTACTCCCTCTGAGAACCCATGAAATGTGACTTCAAGAAATATGATGATG 2220
 OY 2221 ttacactctgttaaaataatcatcgaaaagtgaataatattataggacgaataagaaga 2280
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2221 TTTTACTCTTGTGTTAAATATCATCGAAAGTGAATAATATTATTAAGCGAATATGAAAGG 2280
 OY 2281 aaactctgagaaacaatgctgaaatttgataaaagaacttgaaatttatatacaagaaga 2340
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2281 AAATCTGGAACAAATGCTGAATTTGATTAAGAAAGAACTGAATTTATTAACAAGAAAGAAAA 2340
 OY 2341 cagaattctgtatcaacgggtacagctcttactagaactgttccagactagtgagaaga 2400
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2341 CAGAAATTTGATTTACCCGGTACAGCTTCTTACTAGACTGTTTCAAGACTAGTGAAAGAGA 2400
 OY 2401 atcagaattttgggttttcaacaagctacacagaaaagatgtgatatgtcaatgttttga 2460
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2401 ATCAGAAATTTTGGGTTTACAAAGCTACACAGAAAGAGTGGATATGCAAAATTTTAA 2460
 OY 2461 tatttggagaagaagaatcagataatctgttcaacaggttttctcgtcccccctcaac 2520
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2461 TATTTGGGAAGAGGAATATTCAGATTAATCTGTTAAACAGCGTTTCTCGTCCCTTCAAC 2520

OY 2521 ttcatattactgctctttagaatttaaaaaatgacatacttccagaagtataagat 2580
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2521 TTCTACATTACTGGCTTTTAGAATTTAAAAAATGCATACTTTTGAAGTATAGAT 2580
 OY 2581 catatcttgaatttttataaatatgataatgaatccttagatttttaccagctt 2640
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2581 CATATCTTGAAATTTTATTAATATGATGAAATTTCTAGAGATTTTTCACACTTT 2640
 OY 2641 gtttaagaccccaatgtaataatataataatttgcatttccagaagttaga 2700
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2641 GTTTACAGACCCCAATGTAATATTAATTAATTTTGCATTTTCTACAGAAATTTGA 2700
 OY 2701 taactgtlaaagaataatcagaataaacttgagctgtctgttctacattaaaaa 2760
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2701 TACCTTTTAAAGAAATTTACAGAAATTAACCTGTGTGCTGTTTACATTAATAA 2760
 OY 2761 aaaaaaaaaaactcgag 2780
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 2761 AAAAAAAAAAAACTCGAG 2780

RESULT 2
 AB028070
 LOCUS AB028070 2719 bp mRNA PRI 04-JUN-1999
 DEFINITION Homo sapiens mRNA for activator of S phase Kinase, complete cds.
 ACCESSION AB028070
 VERSION AB028070.1 GI:4996097
 KEYWORDS activator of S phase Kinase.
 SOURCE Homo sapiens Hela cDNA to mRNA, clone:H37.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Kumagai,H., Sato,N., Yamada,M., Mahony,D., Seghezzi,W., Lees,E.,
 Arai,K. and Masai,H.
 TITLE A Novel Growth- and Cell Cycle-Regulated Protein, ASK, Activates
 Human Cdc7-Related Kinase and Is Essential for G1/S transition in
 Mammalian Cells
 JOURNAL Mol. Cell. Biol. (1999) In press
 REFERENCE 2 (bases 1 to 2719)
 AUTHORS Masai,H. and Kumagai,H.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAY-1999) to the DDBJ/EMBL/GenBank databases. Hisao
 Masai, Institute of Medical Science, University of Tokyo,
 Department of Developmental Biology, 4-6-1 Shitokaneda, Minato-ku,
 Tokyo 108-8639, Japan (E-mail:hisao@ims.u-tokyo.ac.jp,
 Tel:81-3-5449-5661, Fax:81-3-5449-5424)
 FEATURES
 location/Qualifiers
 1..2719
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_type="Hela"
 /chromosome="7"
 /clone="H37"
 /map="7q21"
 518..1222
 /gene="ASK/ H37 (hubb4)"
 /gene="ASK/ H37 (hubb4)"
 518..1222
 /note="C-terminally truncated due to alternative splicing;
 an alternative splicing variant of ASK, a regulatory
 subunit for hucdc7 kinase. This protein could bind to
 hucdc7 but may not able to activate its kinase activity."
 /codon_start=1
 /product="activator of S phase Kinase"
 /protein_id="BA078327.1"
 /db_xref="GI:4996098"
 /translation="MNSGAMRIHSGHGGGCTQVYNNKRPISLKTIDNRPESKCK
 PLMGVVFYLDLPSTVISEKLDIDLGGRVEEFLSKDISLISNKKKEKFAQTIGRI
 SPVSPESAVYTAETITSPHSHDGSSFFSPDYCLSGKLIVKRAIKRDFIPISNLS
 NALSWGKTLIHDDIRYITFQKKRELYLTKSSTSVRDGKRVGGAQKTRFRGLRKP
 FVKVDMDSQSPVHLM"
 2563..2569

polyA_site

BASE COUNT 932 a 522 c 607 g 658 t
 ORIGIN
 Query Match 95.3%; Score 2648; DB 9; Length 2719;
 Best Local Similarity 97.8%; Pred. No. 0;
 Matches 2719; Conservative 0; Mismatches 0; Indels 61; Gaps 1;

```

OY 1 aatttcgacagagctctctgaagctcgcccaagacctgaagcggcgagacgagacggcg 60
Db 1 AATTTCGACAGAGCTCTCTGAGGCTCGCCAGACCTGACGCGGACCGAGCCGCGG 60
OY 61 gctcgaagctgaagagagcaagagatgagagcgagggtagagcgagaaacaaacctgcag 120
Db 61 GCTCGAGACTGAGAGACCAAGCATGAGCGGGTAGAGCGGAAACCAACCTGCAG 120
OY 121 ggcacagacgagcgcgagagaaagagacgagctgaagggcgagggcgagagcgagaa 180
Db 121 ggcacagacgagcgcgagagaaagagacgagctgaagggcgagggcgagagcgagaa 180
OY 181 ggcacagacgagcgcgagagcgagagcgagggcgagggcgagggcgagggcgaggg 240
Db 181 ggcacagacgagcgcgagagcgagagcgagggcgagggcgagggcgagggcgaggg 240
OY 241 ggcacagacgagcgcgagagcgagagcgagggcgagggcgagggcgagggcgaggg 300
Db 241 ggcacagacgagcgcgagagcgagagcgagggcgagggcgagggcgagggcgaggg 300
OY 301 ctgcgttgccttgccttcctcctcctcctcctcctcctcctcctcctcctcctcctc 360
Db 301 CTGCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGC 360
OY 361 ccgacgtcgaagcgagcgaactcctcctcctcctcctcctcctcctcctcctcctcct 420
Db 361 CCGACGTCGAGAGCGGCTACCTTACGCTGAGAGCGCTTACCTGAGAGCGGCTTACCTG 420
OY 421 cggcgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 480
Db 421 CGGCGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 480
OY 481 ttttcgagacgagatgtagtgacgagcgagctgcagaaactcgaagacgagagat 540
Db 481 TTTTCGAGACGAGATGTAGTGACGAGCGAGCTGCAGAAACTCGAAGACGAGAGAT 540
OY 541 ccacagtaagagacatcccgaggtggaatcccaatcaaaatgtaaaatgtaaaatgta 600
Db 541 CCACAGTAAGAGACATCCCGAGGTGGAATCCCAATCAAAATGTAATGTAATGTAATG 600
OY 601 tctgaatctctgaataactgaatacagagcagaataatcccaatgtaagcactctgggg 660
Db 601 TCTGAATCTCTGAATAACTGAATAcagagcagaataatcccaatgtaagcactctgggg 660
OY 661 aaaaagtaatttaactgaacttaactctcctcctcctcctcctcctcctcctcctcct 720
Db 661 AAAAGTATTTTACTTGAATTAACCTTCTGTCACCATATCTGAAAACTTCAAGGACAT 720
OY 721 taagagacgtgagggcgaggtggaagaattcctcagcaagaatataagttatctatctc 780
Db 721 TAAAGATCTGAGGGCGAGGTGAAGAAATTTCTACGAAAGATATCACTTATCTTATTT 780
OY 781 aaaaagaagaagagcgaattgacaaacttgaggtgcaattctccgtacaaatgc 840
Db 781 AAAATGAAGAGAAAGCTAAATTTGCAAAACCTTGAGGCGAATTTCTGTAACCAAGTCC 840
OY 841 agaatctgcatactgcgagaacacttaactcctcctcctcctcctcctcctcctcctc 900
Db 841 AGAATCTGCATATCTGCAGAAACACTTCACTCCTCAGCATGATGAGATTCATTT 900
OY 901 taagtccacgacagagtggtttaagcagaggaataattatgattgaaagagctatcaa 960
Db 901 TAAGTCCACGACAGAGTGTGTTTAAAGCAGAGGAAATTAATTAAGTTGAAAAAGCTATCA 960
OY 961 ggaacagatattatctcctcaaatagatataataaatcttgcatagggagataaa 1020
  
```

```

Db 961 GGACCATGATTTTATTTCTTCAATAGTATTTATTCMAATGGCTTGTCATGGAGTAA 1020
OY 1021 aattctcatatgatgacatatactcactatgaacaaagaaagaaagagttgattt 1080
Db 1021 AATTCTCATATGTGATGACATTTAGATATCTACATTTGACAAAGAAAAAGATTGTATT 1080
OY 1081 actcaagaaatcaagtaacttaagtaagagatggggcacaagagttgtagtgatgaca 1140
Db 1081 ACTCAAGAAATCAAGTACTTCAATAGATGAGTGGCAAAAGAGTGTAGTGTGTCACA 1140
OY 1141 aaaaacagagacgagagactcaaaaagcctcttgtaaaagtggaaagatatgagcaact 1200
Db 1141 AAAAACAGAGACGAGAGACTCAAAAAGCCTTTGTAAAGTGGAAAGATATGACCCA--- 1197
OY 1201 ttatagccatttatctcagctgacccaatgcttataaattatctatccagaa 1260
Db 1198 -----AA 1199
OY 1261 gccctgagtcacatttgatgtagacaagcactagatgcaaaagcaaacctcaggttaa 1320
Db 1260 GCCCTGAGTCACTTTGATGTAGACAAACCATAGTATGCAAAAGCAAACTCAGGTTAA 1259
OY 1321 actaagatctcaaacagatggagataagtatggtggaacctcaattcaactcagttgaa 1380
Db 1321 ACTAAGATCTCAAAcagatggcagATAGTATGATGGAACCTCAATTCAACTCAGTTCAA 1319
OY 1381 agagaaagaaagaaagaaatatttgaaatggttgcctgcagaagatatgaagatcaga 1440
Db 1380 AGAGAAAGAAAGAAAGAAATATTGTCAATGTTGCTTCAGAAATATGAAGATCTAGAAC 1379
OY 1441 tcaacctcgaatgtagcaacacagaacattgcaacagatgaacagatatacaagttgtga 1500
Db 1440 TGACCTCTAGTACGACACACAGAACTTTGCAAGATATACCAATATCAAAAGTTGTTGA 1439
OY 1501 tgaattgtaactcaagttgatttgcactctggaatatagaagaaagacacactaaaa 1560
Db 1501 TGATATTTGTATATCAAGTTAGTTTGTGACTTTGTGAATATGAAAGAACACCTTAAATA 1439
OY 1561 gaaaagaataaatacagctgtctgaaaccccttccctcctcctcctcctcctcctcct 1620
Db 1560 GAAAAGATATAATATACAGTGTGTGATCCCTTCTCCTGTTTGTGCAAGTGTCTTAAATA 1559
OY 1621 gactgaacaaagaaagaaagttgaaatlgcaacatacttccagaagaaattgccaagaa 1680
Db 1620 GACTGAACAAAGAAAGAAAGTGAATGCAACATATTTCTCAGAAAGATTTGCCAGAAAGA 1619
OY 1681 tgaatacaacagtgagagagagaaattcctcgtataaagagaccagaagaaactgaa 1740
Db 1680 TGATATCAACAGTGAAGAGCAGAAATTTCTGTATTAAGAGACCAGAAATCGAAAAAAA 1679
OY 1741 gctcgtgttcttctcagaagccatccccaacctcgaatgaattgagaggtctaaaga 1800
Db 1740 GCTCCTGTTTATTTCAAGGCCATGCCCACTTCAAAATGAATGAGGGCCTTAATGA 1739
OY 1801 gaaaatgagcaataatggttccatggttaagtcagcgtgaagatgacatgaacagpat 1860
Db 1800 GAAAATGAGTAAATTAATGTTCAATGTTAAGTACAGCTTGAAGATGACATGAAGCAATTT 1799
OY 1861 tacacagctaccttataataaacaacagaaatgcaattctcgacaattccgaacacac 1920
Db 1860 TACACAGCTACCTTACATTAATAAAACAAACAGAAAGATCTTGTGCAATTTCCGAACAC 1859
OY 1921 attaaagtgaatgaacttagaagaactaaaggttagatcaatataatgataacagagc 1980
Db 1920 ATTAAGTGAATGACTTGAAGAGACTAAGGGTAGATCACTAATAATGTAACTACAGGC 1919
OY 1981 atctgtatgcttctcattctcagtagacagataatagtgatctcaacaaacaagaatc 2040
Db 1980 ATCTGTATGCTTCTCATATTTCAATGATCAATATAGTGTGATCTCAACCAAAACAGAGTC 1979
OY 2041 agatcgtgcttcttccagcaagaattcagaagaaagagccttcatcattctaac 2100
  
```

Db	1980	AGACAGTGTCTTTTTCACGAAAGATTCACAGAAAGAACCTTCATTCATATTTTAC	2039
QY	2101	tcatgatcttgcgttcgtatcaacataaagctltcaagaagacctaactcgtltcaaggcaa	2160
Db	2040	TCAATGATTCCTGGTCTGTATTAACAATTAACAGTTCACAAAGACACCTAATCTTCAGCGAAA	2099
QY	2161	ggctccattccatctctctctctgaaggaaacccaatgaatgtgacttcaagatatgagtag	2220
Db	2100	GGCTCCATTCATTCATCTCTCTCTGAGGAACCCAAAGAAATGTGACTTCAGAAATATGGATAG	2159
QY	2221	tttacccttcgttaaaatacaatcaaaagatgaaataatattatgaacgaatatagaaga	2280
Db	2160	TTTACCTCTTCGTGTAATTAATACATGAAAAGTGAATAATATATGAGACGAAATATGAAAAGA	2219
QY	2281	aaatctggaaccaaactgcgaatttgaataaagaacactgaatttatacacagaagaagaa	2340
Db	2220	AAATCTGGAAACCAATGCTGATTAATTTGATTAAGAAGACTGAATTTTATACAGAAAGAAAA	2279
QY	2341	cagaatttgcgttcaccggtlaacagcttactagacttgcttcagaactgtgaagaaa	2400
Db	2280	CAGAAATTTGATGATTCACCGGTACAGCTTACTATGACTTGTTCAGACTGTGTAAGAGAA	2339
QY	2401	atcgaatcttttgggtgttccaaagctaacacagaagaagatgatatgaagaatgtttaga	2460
Db	2340	ATCGAATTTTGGGTTTTCACAACTTACACAGAAAGAGTATATGCAATGTTTTGA	2399
QY	2461	tattctgggaagaggaataatcagaataatcgttaacagcgttttctccctccctaac	2520
Db	2400	TATTTGGGAAGAGGAATAATTCAGATTAATCTGTATACAGCGTTTTCCTGTCGCCCTTCAC	2459
QY	2521	ttctacatttacttgcgttcttgaagaattaaanaatgcatacttccagaagtgaagaagat	2580
Db	2460	TTCTACATTTTCTGCTGCTTTTGAAGATTTAAAAAATGCATCTTTCAAGAAAGTAAAGAT	2519
QY	2581	catcttcttgaatcttltataaataatgatatggaatctttagaattttttaccagcttc	2640
Db	2520	CATATCTTGTAATTTTATATAATATATATGAGAAATTTCTTAGATTTTATACACAGCTTT	2579
QY	2641	gtttacagaccacaatgtaaatatataaataatcatttgcatacttctcagaagatgaa	2700
Db	2580	GTTTACAGACCCAAATGTAAATATTAAAAATTAATTTTGCATTTCTACAGAAATTTGAA	2639
QY	2701	tacctgtttaaagaanaattlacagaataaacttgcactgtgctgttctgtttacatataaaa	2760
Db	2640	TACCTGTAAAGAAAATTTACAGATTAACCTGTGACTGTGCTGTTCATATTAATAAAA	2699
QY	2761	aaaaaaaaaaaaaactcgag	2780
Db	2700	AAAAAAAAAAAAAAAAAAGCTCGAG	2719
RESULT	3		
LOCUS	AF160249		
DEFINITION	Homo sapiens 2474 bp mRNA	PRI	01-JUL-1999
ACCESSION	AF160249		
VERSION	AF160249.1	GI:5306124	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mommalai, Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	(bases 1 to 2474) Jiang, W. and Hunter, T.		
JOURNAL	Mammalian Cdc7/Dbp4 Protein Kinase Complex is Essential for		
REFERENCE	Initiation of DNA Replication		
AUTHORS	Unpublished		
TITLE	2 (bases 1 to 2474)		
JOURNAL	Jiang, W. and Hunter, T.		
REFERENCE	Submitted (17-JUN-1999) Molecular Biology and Virology Laboratory		
AUTHORS	The Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA		
TITLE	92037, USA		

FEATURES		location/Qualifiers
source		1..2474
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
CDS		217..2241
		/note="Hscd/c7 protein kinase regulatory subunit"
		/codon_start=1
		/product="Dbf4 protein"
		/protein_id="Aad41911.1"
		/db_xref="GI:5306125"
		/translation="MNSGAMRIHSKGFPGQIVKNENRSLKSLTDNRPEPLNGKVIYLDLPSVTLSERKLQDKIDKGVEEFLSKDLSYLISNNKEARFSPVSPSPSATTAATTSRPHSDSSFSKSPTYVLCSGLIKLVKALTDHPFISSNALSWGVKLIHRPDIIRYLEOKRKELLYLKSSSTVRDGSRKGSQAQRTTRIGFVVEDMSQLRPDYFDLTNMPFINYSIQKPSPEVPVDAPSSMKOCTQVLRIGKYSGTSLDLKKRRKKGYCECLOKYEDELFTLLSEORINAOSONOVDDVPFEVEEKDTPKKRIRKRYSVGLSPVASLYLKTKQKRVLETOLISQKCOOEKEONLEKTEOTETEKILPISEPPIPPSNRLRIANRSMKSMTSTARDIIRLPJHKRNQCETLIDISEHTLSENDLELRVDHRCNTVQASVHSVSDSTDSGSO DYVALPKDLKEDLHSIFTHDSGLITINSQEHNLVAQAKAPHHPPEEPNCCDSLDPGKIHRKVKITLGRNRKEMLEPNAEKDKEFERITQENNICSSPOSUILLSEEKSEFLGPTSTSEKSGICNVLDIWEENSDDLPAFPSPSPSTSFPTGF"
BASE COUNT	888 a	449 c 480 g 657 t
ORIGIN		
Query Match	88.4%;	Score 2458.8; DB 9; Length 2474;
Best Local Similarity	99.9%;	Pred. No. 0;
Matches 2471:	Conservative	0; Mismatches 2; Indels 1; Gaps
Oy	301	cctcgttgccttcgctcgcccttccttcgcgccttcggagccggatccggcccggaaac 360
Db	1	CTCGTTTGCTTGCG-CTTCCTCTCCGGCCTTGAGACGGATCCGCCCGGAAC 59
Oy	361	ccgacctgcaaacggcgtaactcttaactgcgtagaaggccgtagctcgcggaagaagagag 420
Db	60	CCGACTCTGCACAACGGGTACCCTACCTCCGTAAGAGCGCGTAGCTGGGGGAAGGAGAAGG 119
Oy	421	cggccgtccttcaacaaggccgggggaagccgctcttcgcgagctgccggttggaac 480
Db	120	CGGCGTCTCTCAACAGCGCGGGGGAAGCCGTCTTCCGGCGCTGCCCGTGGACACT 179
Oy	481	tcttcggaccacgatgtagtgcgcggcgagctgcatagaactccgagaccatgagat 540
Db	180	TTCCTCGGACCCAGCATGTAGTGTCGGGGGACGTGCCATGAACTCCGAGACATGAAGAT 239
Oy	541	ccaagtaagaagacattccagggatgaattccaagtccaataatgaaaaaaacagacatc 600
Db	240	CCACAGTAAAGACATTTCACGGGTGCATCCAAGTCAAAAATMAAAAAACACACATC 299
Oy	601	tctgaatctctgaanaactgatacaagccagaanaatccaatgtaagccaactttggg 660
Db	300	TCTGAATCTCTGAAAATCTGATAACAGGCCGAAAAATCCAAATGTATAGCCACTTTTGGG 355
Oy	661	aaaagtatttaaccttgacttaccttcgyltaccatalctlgaaaacctcaaaaagacat 720
Db	360	AAAAGTATTTTACCTTGACTTACCTTCTGTCAACATATCTGAANAACCTCAAAAGGACAT 419
Oy	721	taaggatctggaggggcgagttitgaagaatttctcaagaagaatacaagtatatcttattc 780
Db	420	TAAAGATCTGGAGGGCCAGTGAAGAAATTCACGCAAAATATTCAGTTATCTTATTTC 479
Oy	781	aatlangaagaagaactltgtaccaaaccttggatcgaaattctctctgtaccagcttc 840
Db	480	AATATGAAGAAGAGCTAATTTTGCACAAACCTTTGGGTGCAATTTTCTCTGTACCAAGTCC 539
Oy	841	agaatctgatatactcgagaacaacattcaactcattccaccgacatgltggaattcat 900
Db	540	AGAATCTGATATACTGCAGAACCACTTCACTCACTCCATCCGACATGATGAAGTTCATT 599
Oy	901	taagtaaccaagacacatgctgtttaagcagaagaaattatagttgaaaaagcatacaa 960
Db	600	TAAAGCACCAACACAGTGTTTTAAAGCAGGGAATTTATAGTTTAAAAAGCTATCAA 659

Moore Drive, Research Triangle Park, NC 27612, USA

FEATURES
source

1. 2460
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7g21"

1. 2460

/gene="DBF4"

/gene="DBF4"

200. 2224

/codon_start=1

/product="DBF4-like protein"

/protein_id="A4537.1"

/db_xref="GI:556281"

/translation="MNSGAMRHSKHFGOGIIVKRNKRLSLKLTQDNREPKSKC
PLMGKVFYLDLPSVTISEKLQIDKLGSRVEEFLSKDYSYISMKKEKFKQTLGRI
SPVSPESAYTAETTSPPSHDGSFSPDVCISRGKLLVKAIKDHPFIPSNILS
NALSMGVKLIHIDDIROYTEOKKRLYLKSTSPVDSGKRVGSAOKTPRGRLKP
FKVYEDMSQLRREYLOLTNNPFTYSTOKPSPPVDKPSMOKOTYKRLAIDQD
KYGGTSIOQLKEKKKGYCECCLOKTYDLEHLISEDRNFAGSNQYVDDIVSKL
VEDVEYEKDTPPKRRIRYSGVSLPVSASVLEKTEOKKEVLAHI SQDCQEDDTV
KQNFLEYKQTEKKLLFISEPIHPSENLKGLNEMKSNKCSMLSTADDIROMETQ
LPLKNGKQECILDISEHTLSEMDLELRVDHYKCNIOASVHSDSTDNSSGSPORS
DTVLFPARKDIKEMDHSITPDHSGLITINSQEHLYVAKAPPHMPEPNPCDKNM
DSLPSGKHRRKYYKTIILGRNKNENLEPNAEPDKRTFTIOENRITSSPQSLDLDFOT
SEKSEFJGFTSYTEKSGICNVLDIWEENSDNLTAFFSSPSTFTGF"

BASE COUNT 891 a 445 c 475 g 649 t

ORIGIN

Query Match 88.3%, Score 2456, DB 9, Length 2460;

Best Local Similarity 100.0%, Pred. No. 0;

Matches 2456; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 cttctctcccgccgttgagccgcatccgcccgaaccgaccctgcaagcagcgt 378
Db 1 cttccctcccgccgttgagccgcatccgcccgaaccgaccctgcaagcagcgt 60
QY 379 acctactcgttagagccgctagctgcggaagagagagcgccgtctctgcaag 438
Db 61 acctactcgttagagccgctagctgcggaagagagagcgccgtctctgcaag 120
QY 439 gccgggggaagccgtgcttcgagctgcccgttgagacacttcccgagccagcatg 488
Db 121 gccgggggaagccgtgcttcgagctgcccgttgagacacttcccgagccagcatg 180
QY 499 taggtccggcgagctgcatgacatccggaagcatgagatccacaagtaagacatt 558
Db 181 taggtccggcgagctgcatgacatccggaagcatgagatccacaagtaagacatt 240
QY 559 ccagggtggaatccaaatgcaaaaaaagacacacatctcgaatctctgaaac 618
Db 241 ccagggtggaatccaaatgcaaaaaaagacacacatctcgaatctctgaaac 300
QY 619 tgataacagggccagaataatccaaatgtaagcacttgggggaaagtaatttaccttga 678
Db 301 tgataacagggccagaataatccaaatgtaagcacttgggggaaagtaatttaccttga 360
QY 679 cttacacttgtaacatactcgaataactcgaagaagacatgaagatctgggaagggc 738
Db 361 cttacacttgtaacatactcgaataactcgaagaagacatgaagatctgggaagggc 420
QY 739 agttgaagaattctcagaagaataatcagttatcttattccaataagaagaagctaa 788
Db 421 agttgaagaattctcagaagaataatcagttatcttattccaataagaagaagctaa 480
QY 799 atttgaacaacttggtgcgaattctcctgtaacaaagtcacgaatctgcatatctgc 858
Db 481 atttgaacaacttggtgcgaattctcctgtaacaaagtcacgaatctgcatatctgc 540
QY 859 agaaacacttcaactccagcatgtagaagttcattaaatcaccagacagat 918
Db 1621 agaaacacttcaactccagcatgtagaagttcattaaatcaccagacagat 1680

Db 541 AGAAGCACCTTACCTCATCCAGCCATGATGGAAGTTCATTTAAGTCACACACACAGT 600
QY 919 ggtttaagcagaggaataatattagttgaaaaagctatcaagacatgattatctcc 978
Db 601 ggtttaagcagaggaataatattagttgaaaaagctatcaagacatgattatctcc 660
QY 979 ttcaaatagatatttcaaatcccttgtaatgggagtaaaaattcttcaattatga 1038
Db 661 ttcaaatagatatttcaaatcccttgtaatgggagtaaaaattcttcaattatga 720
QY 1039 catatagatatttcaaatcccttgtaatgggagtaaaaattcttcaattatga 1098
Db 721 catatagatatttcaaatcccttgtaatgggagtaaaaattcttcaattatga 780
QY 1099 ttcaaatagatatttcaaatcccttgtaatgggagtaaaaattcttcaattatga 1158
Db 781 ttcaaatagatatttcaaatcccttgtaatgggagtaaaaattcttcaattatga 840
QY 1159 actcaaaaagcccttgtaaggtggaagatgagcccaacttataagccattatct 1218
Db 841 actcaaaaagcccttgtaaggtggaagatgagcccaacttataagccattatct 900
QY 1219 tcaagctgaacataatgaccttataataatttcaatcaagaagccctgagctcattga 1278
Db 901 tcaagctgaacataatgaccttataataatttcaatcaagaagccctgagctcattga 960
QY 1279 tgaagaagaacatcagatgtaagaagaacacaggttaaaactaagaatccaacaga 1338
Db 961 tgaagaagaacatcagatgtaagaagaacacaggttaaaactaagaatccaacaga 1020
QY 1339 tggcgataagatggtggtggaaccccaatccaatccaggttgaagaaggaagaagaag 1398
Db 1021 tggcgataagatggtggtggaaccccaatccaatccaggttgaagaaggaagaagaag 1080
QY 1399 atatttgaatgttgccttgtaagaataatgtaagaacccaacttcaagtgagca 1458
Db 1081 atatttgaatgttgccttgtaagaataatgtaagaacccaacttcaagtgagca 1140
QY 1459 acaagaagaacttgcagagatgtaacagatcaagttgtatgatatgtatctaagtt 1518
Db 1141 acaagaagaacttgcagagatgtaacagatcaagttgtatgatatgtatctaagtt 1200
QY 1519 agttttgacttggtaataatgaaagagacacacccaataaagaagaataaataacag 1578
Db 1201 agttttgacttggtaataatgaaagagacacacccaataaagaagaataaataacag 1260
QY 1579 tgttgaatcccttctccttcttctgcaagtgcttcgttaaaagacatgaacaaagaaaa 1638
Db 1261 tgttgaatcccttctccttcttctgcaagtgcttcgttaaaagacatgaacaaagaaaa 1320
QY 1639 agtgaattgcaacataatctcagaagaatgccaagaagatgatacaacagtgaaaga 1698
Db 1321 agtgaattgcaacataatctcagaagaatgccaagaagatgatacaacagtgaaaga 1380
QY 1699 gcaagaattctctgataaagaagccaggaactgaaaaaagctccgttcttcaaga 1758
Db 1381 gcaagaattctctgataaagaagccaggaactgaaaaaagctccgttcttcaaga 1440
QY 1759 gccacatcccccaccccttcaaatgaatgagaagggcttaatgagaataatgtaataatg 1818
Db 1441 gccacatcccccaccccttcaaatgaatgagaagggcttaatgagaataatgtaataatg 1500
QY 1819 ttccatgttaagtaagctggaagatgacataagaacagaattttaacagctaccctaca 1878
Db 1501 ttccatgttaagtaagctggaagatgacataagaacagaattttaacagctaccctaca 1560
QY 1879 taaaacaacaagaatgagcttcttgaacatttcggaacacacattgaatgaatgaatga 1938
Db 1561 taaaacaacaagaatgagcttcttgaacatttcggaacacacattgaatgaatgaatga 1620
QY 1939 agaaagaactaaggtagatcaactataaatgtaacatacagacatctgtaatgttctga 1998
Db 1621 agaaagaactaaggtagatcaactataaatgtaacatacagacatctgtaatgttctga 1680

QY	1999	tttcagtaacgaataagtgagcatctaacccaaacagaagtcagatctcgtctttcc	2058
Db	1681	TTTCAGTACAGATAATATGTGATCTCCACCAAAACAGAAATCAGATACGTGCTTTTCC	17400
QY	2059	agcaaaagatctcagaagaaaggactctatctcaattctcctaattctcgtctcat	2118
Db	1741	AGCAAAAGATCTCAGAGAAAGACCTTCATTTCAATATTTCATCATATTTCTGCTCAT	18000
QY	2119	aaacaataacagttccacaagaagcaaccctaacctgtctcaggcaaaagctccattccatctcc	2178
Db	1801	AAACAATAAACAGTTCCACAGAGCACTTCATTTCAATATTTCATCATATTTCTGCTCAT	18600
QY	2179	tccttgagaaacccaatgaaatgtgacttccaagataatgataatgtttaaactctgtitaaat	2238
Db	1861	TCCTGAGGAACCCCAATGAATGTGACTTCAGAAATATGGATGTGTTACCTTCTGGTAAAT	19200
QY	2239	acatcgaaaagtgaanaataatctaggaacgaatatgaaaagaanaatcttgaaacccaatgc	2298
Db	1921	ACATCGAAAAGTGAAAATTAATATTACGAGCAATATGAAAAGAAAATCTGGAACCAAAATGC	19800
QY	2299	tgaattctgataaagaacctaatcttctacacaagaagaagaacgaatttgatgcttacc	2358
Db	1981	TGAATTTGATTAAGAACCTGAAATTTATTTACACAGAACAAAACGAAATTTGTATGTTGCTC	20400
QY	2359	ggtacagctcttctacgaactgtgttgaagtagtgaaagagaatccgaatttttggctt	2418
Db	2041	GGTACAGCTTTTACTAGCTGCTGTTTCAGCTATGATAGAGAAACACAAATTTTGGGCTT	21000
QY	2419	cacaagctacacagaaagaagtcggtcatatgcagtgcttttagatacttggagaagagaaa	2478
Db	2101	CACAAGCTACACAGAAAGAGTGTATATGCATAATGTTTAAATTTTGGAGAACAGAAA	21600
QY	2479	ttcagataactcgttcaacagcgttttctccgctcccttcaacttctacatttacttgcgtt	2538
Db	2161	TTTCAGATATCTGTTTAACAGCGTTTTCGTCGCCCTTCAACTCTACATTTACGTGCTT	22200
QY	2539	ctcgaattcaaaaaacgcatacttctcagaagtgaataagatactcttgaattctt	2598
Db	2221	TTTGAAATTTAAAAAATGCATATCTTTTCAGAACTGATTAAGATCATATTTCTGAAATTTT	22800
QY	2599	ataaataatcttgaaaactctcttgagaatttcttaccagcttgttccagaaccgaatgct	2658
Db	2281	ATAAATATGTATGGAATTCCTTATGGAATTTTATTTTACCACCTTGTTTACAGACCAAAATG	23400
QY	2659	aaattctttaaataaataattctgaattctctcagaatttgataccctgttaagaagaaat	2718
Db	2341	AAATATTTAAAAATTAATATTTTCCAAATTTTCTACAAATTTGATATCTGTTTAAAGAAAAT	24000
QY	2719	tacagaataaacctgtgtactgctgtctgttttcaatcaataaaaaaataaaaaa	2774
Db	2401	TACAGAAATAACTGTGACTGCTGTCTGTTTACATTTAAAAAATAAAAAAATAAAAAA	2456
RESULT	5		
ACOL18977/c			
LOCUS	ACOL18977	236236 bp	DNA
DEFINITION	Homo sapiens chromosome 10 clone RP11-386C23, WORKING DRAFT		26-JUL-2001
SEQUENCE	SEQUENCE, 12 unordered pieces.		
ACCESSION	ACOL18977		
VERSION	ACOL18977.6	GI:1502187	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEPIN.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 236236)		
TITLE	Smith,D.R.		
JOURNAL	Genome Therapeutics Corporation Sequencing Center: Human Genome		
REFERENCE	Unpublished		
AUTHORS	Sequence Data		
	2 (bases 1 to 236236)		
	Smith,D.R.		

```

TITLE      Direct Submission
JOURNAL    Submitted (25-DEC-1999) Genome Therapeutics Corporation, 100 Beaver
           Street, Waltham, MA 02453, USA
COMMENT    On Jul 26, 2001 this sequence version replaced g1:14091812.
           -----
           Genome Center
           Center: Genome Therapeutics Corporation
           Center code: GTC
           Web site: http://www.genomecorp.com/
           Contact: gtc-seqcenter@genomecorp.com
           -----
           Project Information
           Center project name: hg092
           -----
           Summary Statistics
           Sequencing vector: N/A
           Chemistry: dye-terminator Big Dye; 100% of reads
           Assembly program: Phrap; version 990315
           Consensus quality: 228735 bases at least Q40
           Consensus quality: 230209 bases at least Q30
           Consensus quality: 231169 bases at least Q20
           Insert size: 235235; sum-of-contigs
           Quality coverage: 6.3x in Q20 bases; sum-of-contigs
           -----
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 12 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *
           1      1146: contig of 1146 bp in length
           *      1147      1246: gap of unknown length
           *      1247      2466: contig of 1220 bp in length
           *      2467      2566: gap of unknown length
           *      2567      4025: contig of 1459 bp in length
           *      4026      4125: gap of unknown length
           *      4126      5165: contig of 1040 bp in length
           *      5166      5265: gap of unknown length
           *      5266      6309: contig of 1044 bp in length
           *      6310      6409: gap of unknown length
           *      6410      7552: contig of 1143 bp in length
           *      7553      8116: contig of 1064 bp in length
           *      8117      8616: gap of unknown length
           *      8617      10307: contig of 1451 bp in length
           *      10308      10407: gap of unknown length
           *      10409      11913: contig of 1506 bp in length
           *      11914      12013: gap of unknown length
           *      12014      13633: contig of 1350 bp in length
           *      13634      13463: gap of unknown length
           *      13464      106707: contig of 93244 bp in length
           *      106708      106807: gap of unknown length
           *      106808      236636: contig of 129429 bp in length.
           *
           Location/Qualifiers
           1. 236636
           /organism="Homo sapiens"
           /db_xref="taxon:9606"
           /chromosome="10"
           /clone="RP11-386C23"
           /clone_id="RPC1-11"
           1. 1146
           /note="assembly_name:Contig7"
           1247. 2466
           /note="assembly_name:Contig11"
           2567. 4025
           /note="assembly_name:Contig12"
           4126. 5165
           /note="assembly_name:Contig13"
           5266. 6309
           /note="assembly_name:Contig16"
           6410. 7552
           /note="assembly_name:Contig17"
           7653. 8716
           /note="assembly_name:Contig18"

```


	misc_feature	8817..10307	/note="assembly_name:Contig20"	
	misc_feature	10408..11913	/note="assembly_name:Contig21"	
	misc_feature	12014..13363	/note="assembly_name:Contig22"	
	misc_feature	13464..106707	/note="assembly_name:Contig23"	
	misc_feature	clone_end:"T7"		
	misc_feature	106808..236236	/note="assembly_name:Contig24"	
	misc_feature	clone_end:"SP6"		
BASE COUNT	70014	a 44567 c 46458 g 74094 t 1103	others	
ORIGIN				
Query Match	Best Local Similarity	76.2%; Score 218.6; DB 2; Length 236236;		
Matches 2329; Conservative	92.2%; Pred. No. 0;	Mismatches 69; Indels 129; Gaps 4;		
OY	254	gccgcgcccgcgttgcagcgttttcaaatcttaaccgcgcgcacccactcgttgtct	313	
Db	218540	GCCCGGGCCGGGTGACGCATTTTCAAACTTCAACCCTGCAGCCCCCCTCGTTTGATT	218481	
OY	314	tgcgccttcctcctccgcgccttgtagcccgatccgcgcgcggaaaccgcactgcagac	373	
Db	218480	TGCGGCTTCTCTCTCCGTGCGCTTGAGCGCATCCGGCCCGGAACCGACCCACAGAC	218421	
OY	374	gcgcgaacctctaactcgtagaaggcsgtgcgtcgcggaaggagaagagcgcgctctgc	433	
Db	218420	GCGGTACCTCTAAGCTGTAAGGCCGTGACCGCGGAAGTAAGAAGCGCGCTCTGTC	218361	
OY	434	aacacgccccgggaaagccgltgcttcgcgcgcgtccgcgggtgcgcacaattctccggagcca	493	
Db	218360	AACAGACGCGGGGAAGCCMTTGGCTTTCGGGGCGCCCGGGCGCACCTTCTCCGGACCCA	218301	
OY	494	gcactgaagtgcgcggcggaactgcgatgaactcccggaagccatagaatccacagtataagg	553	
Db	218300	GCATGTAGGTGCCGGCGGCACCTCCATGAACTCTGGAGCGGTAGAGATCCACAGTAAAGGA	218241	
OY	554	cattcccgaggcgaggaatccgaatcaaaaatgaaaaaacagaacacatctctgaactctcy	613	
Db	218240	CATTTCCAGGGTGGATCCAACTCAAATAATGANAATAAACAGCACATCTTGTGAAATCTCTG	218181	
OY	614	aaaactgtaaacagagccagaaaaatcccaaatgtlaagccactcttgaggaaagtatttac	673	
Db	218180	AAAACGTGATTAACAGGCCAGAAAAAATCCAAATTTAAGCCACTTTGGGGAAAAATATTTCAC	218121	
OY	674	cttgacttaacctctgcgcacacatcltcgaaaaactccaagaagacattaagatctbgga	733	
Db	218120	CTTGTACTTAACCTTCTGTGACCACTATCTTGAAAAACTGCAAAAGACATTAAAGATCTGGGA	218061	
OY	734	ggcggaaglttgaagaattlctcagcaagaatatcacgttatcttatcttcaataagaagaa	793	
Db	218060	GGGCGAGTTGAAGAAATTTCTCGCAAGAAAGTAACAGTTACTTTCAATTAATAAAAGGAA	218001	
OY	794	gctaaattgcacaacacttggctgcgaattcttcctgcgtacccaagtcacagaactcgcata	853	
Db	218000	GCTAAATTTCGCAACAACCTTGGGTGGAATTTTCTCTGTACCAAGTCAGAAATCTGCAAT	217941	
OY	854	actgcagaacaacctcatcactcatcccgaccatgatlgaagatcatttaagtccaccagac	913	
Db	217940	ACTGCAGAAACACACTTCACCTCATTCGCCAGCCTAGTAGCAAGTTCAATTTAAGTCACCAAGAC	217881	
OY	914	acagtgctgtttaagcagaggaaatlatatglttgaaaaagactcaaggacatgattc	973	
Db	217880	ACAGTGCTGTTPAACGACAGGAAATTAATTAGTTGAAAAAGCTAATCAAGACATGATTTTT	217821	
OY	974	attctctaaatagatatattcaaatgcctgtcatagtggaggtaaaaattcttcataatt	1033	
Db	217820	ATTCTCTCATATAGTAATTTATCAAAATGCTTGTGTATGGGGGTAAAAATTTCTCATATT	217761	
OY	1034	gattgacattagatctactacaattgacaacaaaagaagaagcttlattactcaagaata	1093	

Db 217760	GATGACATTAGATATACATTGAAACAAAAGAAAAA - AGTTGCAATTCACGACGAATTC	21770303
QY 1094	agtaactcaagaaagatcggggcaaaaagatctgtagtggctgcacaaaacaagaaca	1153
Db 217702	AGTACTTCAGTAAAGAGATGGGGGCAAAAAGACTTGTAAGTGTGCTGCACAAAAACAGAC	2176433
QY 1154	ggaagaactcaaaaagccttctgtaaaagtggaaatatgagccaacttataagccatt	1213
Db 217642	GGMAAGCTCAAAAAGCCCTTTGTAAAGGTGGMAATGTGAGCCAACTTTPTAGGCCATT	21758303
QY 1214	tatcttcagctcgacaaatctgcttttataatatctctcttcaagaagccctgagfca	1273
Db 217582	TATCTTCAGCTGACCAAAATATCCCTTTATAAATTTCTATACGAAGCCCTGCAGTCCA	21752303
QY 1274	tttgatctgacaagcccactccagtatgcgaagaacaactcaagttcaactcaagatccaa	1333
Db 217522	TTTTGATGTAAGCAACCAACCTATAGTATGCCAAAAGCAACTCAGGTTAAACTTAAGATCCAA	21746303
QY 1334	acacatggcagataagtalgtgtgagacccaattcaactccagttgaaagaagaagaa	1393
Db 217462	ACAGATGGCGSATAAGTATGTGTGAAACTCAATTCACGCTGTAAGAAGAGAAAGAA	21740303
QY 1394	aaagatatctgtgaatgtgtctgtgcagaataatgaagatctagaactccacttaagt	1453
Db 217402	AAAGATATTTTGGAATGTTGCTTCCAGAAATATCAAGATCTAGAAACT - - - - -	21735303
QY 1454	ggagcaaacagaaacttgcacagagtaaccagatcaagttgtgatatgtatct	1513
Db 217354	-----	21735303
QY 1514	aagttagttttagactcttggaatagatagaanaagcacacacttaaaaaagaataaa	1573
Db 217354	-----	21734803
QY 1574	tacagtggttgatccctcttccctgttctctgcaagtgctccgaaagaagctgacaaaag	1633
Db 217347	TACAGTGTGGTATCCCTTCTCCCTGTTTCTCAAGTGTCTGAAAAAGCTGAACGAAAG	21728803
QY 1634	gaaagaatggaatctgacaacatatctctcagaagaatctgcagagagatgatalcaacagt	1693
Db 217287	GAAAAAGTGGATTCGCAACATATTTCTCGAANAAGATTCGCCGGAAGATGATTAACGGTG	21722803
QY 1694	aagagcagaatcttcgtlataaagagcccaaggaactgaaaaaaagctctgttatt	1753
Db 217227	ATAGAGCAGAAATTTCTGTATTAATAAGACCCAGGAAACTGAAAAAA - CTTCCTGTTATT	21716903
QY 1754	tcaagagcccatcccccactccaatgaatctgagaagggcttaatgagaaaatgaagtaat	1813
Db 217168	TCAAGGCCATTCGCCACCCCTTCAATATCACTGAGAGGGCTTAAACAAGAAATACAGTAAT	21710903
QY 1814	aaatcttcagtgtaagtacagctggaagaagaactaaagaacgaatttcaacagactaact	1873
Db 217108	AAATGTTCCATGTTAAGTACAGCTGAAATATACATATACAGAAATTTTACACACTACTCT	21704903
QY 1874	ctacataaaaaaacaagaagatgcattcttgacattccgagacacacattaagaagaaat	1933
Db 217048	CTACATTAATAAAACAGAGATGCTTTTGAACATTTCCGAGACACACATTAAGAGAAAT	21698803
QY 1934	gaactagaagaactaaggtagatctactataaatgtaacatagacatagcactgtacatgt	1993
Db 216988	GACTTACAAAGACTTAAGATAGATTAATCTAATAATGTAACTATACAGGACTCTGTTACAGTT	21692903
QY 1994	tctgattcagtaagataaataagtgatctccaaccaaaaagaagatcaagatctgtgct	2053
Db 216928	TCTGATTTCTGTACAGATTAATATGTGCATCTACACAAAACAAAAGTACAGATCCGTGCTT	21686903
QY 2054	tttccagcaaaagatctcaagaaagaagccttcactcaactattactcaatgattcgt	2113
Db 216868	TTTCCAGCAAAAGATCTCAAGGAAAAGSACTTCAATTCATTAATTTACTCATGATTCGGT	21680903
QY 2114	ctgataacaataaacaagtctcaagaagaccataactgttcaagcaaaagctccactcat	2173

	D	216808	CTGTAACAAATTAACAGTTCCACAAGACCACTAACGTTCAGGCACAAAGCGTCAATCACAAT	216749
OY	2174	actcctctcgaggagaccgaatgaaatgtacctaataaatatgatatttacccttcygt	2233	
D	216748	ACTCTCTCTGGAGACCACATATATGTATACATCAAGAATATGTGATATGATTACCTTGCGT	216689	
OY	2234	aaaatacatcgaaagtgaataatatatttagcagaataatagaaaaaacatctggacca	2293	
D	216688	AAAATACATCGAATAAGCGAAATATTTATTACGACAAATATGAAGAAKAAATCTGSAACCA	216629	
OY	2294	aatgcgtgaatttgataaagaacatgaatttatcacacaagaagaacaacagaaatttgtagt	2353	
D	216628	AATGCTGTAATTTGATTAAGAAGCTAATTTATATACACAGAAAGAAATGATATTTGTAGA	216569	
OY	2354	tcaacggtacagctcttacctagacctgtttccagaactagtagaagaagaatcatgaatttlg	2413	
D	216568	TCAACCGGTACAGTCTTTTACCAGACTTGTTCACACTAGTGAAGACGAATCAGAAATTTTTG	216509	
OY	2414	ggtttccaacaactccacaagaagaagtaglalatgcaabtgltttatgatatatttggagaag	2473	
D	216508	GGTTTTCACAACTTACACAAAAAGATGGTATATGCAATGTTTAAATATTTTGGGAAG	216449	
OY	2474	gaaaatcatgataatcgtgttaaagaaggttttctccgccctccaactctcatacttaact	2533	
D	216448	GAAATTCAGATATCTGTTAAACAGGTTTTCTCGGCCCTTCAACTCTCATATTACT	216389	
OY	2534	ggccttttagaatttaaaaaatgatatactttccagaagaatgataagatactatcttggaa	2593	
D	216388	GCGCTTTAGAATTTAAAAATGATACCTTTTTCAGAGAGATGAAGATCATATTCTTGAA	216329	
OY	2594	ttttataaatatgatagaaatctcttggaa- ttttttaacagcttgattttaacagacc	2652	
D	216328	TTTTTTTATATATGTATACGAAATCTCTAGCATTTTTTTTACGACGTTGTTTACGAGACC	216269	
OY	2653	aaatgtaaattttaaataaataatttgaattttcttcaagaatttgaatacctgtttaa	2712	
D	216268	AAATGTAAATATTTAAATAAATGTTTGCATTTTCTACAGAAATGAATGCCGTGTATAG	216209	
OY	2713	aaaaattacagaataaactgtgtacgtgtctgtgttttaacctttaaataaaaaaaaaa	2772	
D	216208	AAAAATTTACAGATAAATCTGTGACTGTCITATTTTACAGTTTAAAAAAGAAACAAATTT	216149	
OY	2773	aactcga 2779		
D	216148	GACACAA 216142		
RESULT	6			
LOCUS	AF292400	2458 bp	mRNA	
DEFINITION	Cricetulus griseus DBP4/ASK mRNA, complete cds.			
VERSION	AF292400.1			
KEYWORDS	GI:13398340			
SOURCE ORGANISM	Chinese hamster. Cricetus griseus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae; Cricetulus.			
REFERENCE	1 (bases 1 to 2458)			
AUTHORS	Guo,B. and Lee,H.			
JOURNAL TITLE	Cloning and characterization of Chinese hamster homologue of yeast			
MEDLINE	DBP4 (chdbp4)			
REFERENCE	gene 264 (2). 249-256 (2001)			
AUTHORS	Lee,H. and Guo,B.			
JOURNAL TITLE	2 (bases 1 to 2458)			
FEATURES	Direct Submission			
SOURCE	Submitted (02-AUG-2000) Research, Northeastern Ontario Regional Cancer Centre, 41 Ramsey Lake Road, Sudbury, ON P3B 5U1, Canada			
LOCATION/Qualifiers	Location/Qualifiers			
source	I. 2458			
	/organism="Cricetulus griseus"			

CDs

195. .xref=taxon:10029-
/codon_start=1
/product="DbpA/Ask"
/protein_id="AAK21856.1"
/db_xref="gi:13398341"
/translation="MNSGAMRHSHKQHFQGIQVNRKNRPSLKSLKTDNRPKSKCK
PLMEKFYLDPSVITISEKLDKIDKLOGRVEERLSIDLYVSNKREKRAQTLGOV
SPVSPESATYAEETSPHPHSDGSEKSPDSEKSLKSTSVSDSKKAGTTIDKIFRANLSLS
NALTWGVKLTIDDIRYIEOKKKECLSLKSTVSDSKKAGTTIDKIFRANLSLS
FLKDEVNYSRPPYLQLTSPINATATPCSPDIEKPSVNOAKQVDPPTDSD
KCGPFDVOLKREKRRKGYECCLQKEDYLEELHSEKHNKFNQSOAYVDVLSKL
VFEPFREDIDPKRRKRYKRVLSVSESVSNVYKTEREKQLALELFOKDWESKQGL
PERFEDIDCEDIQKDEORLVLASEPSVSTGLKROEKASLMEFQDWESKQGL
TDLPKNSANDLQKEDLHPFVGHSDTLNTSKEDQITVAKTGPCGQOQNEDETER
MTNPLDCKTIORRYKLLGKKKXNDVPSLEIDKRTFELMCEDRCSFVSLDLFLF
QTSGSPLEHRTYTESNSGLCDVLDYMDENSSLSLSTFESSPSASTIGP"

BASE COUNT 807 a 467 c 533 g 651 t

ORIGIN

Query Match 55.3%; Score 1537; DB 10; Length 2458;
Best Local Similarity 79.4%; Pred. No. 7.9e+282;
Matches 1963; Conservative 0; Mismatches 453; Indels 53; Gaps 10;

QY 330 ggcgccttgagccggatcgcggcccggaaccgcgaacctgcagacgcggtacctactgc 389
Db 8 GCCCCTTGAGCCCGGATCCGCCGCCCGAACCACCTCGAGACGGGGAGACTCTACGCG 67
QY 390 gtagaagccgtagctgcggaagagagagagcgcgctccctgtcaacagccggcggaag 449
Db 68 GTAGAGGCGCGTAGCTGGCGGAGAGAGAGGGCGCGCTCTGTGTAACAGG -CGGGGAGAG 126
QY 450 ccgtgcttcgcggctgcccgctgcacaccttctccggaccacagatgtagtgcgg 509
Db 127 CCCTGCTTCCGGCGCTGCCGGTGGCGACACTTCTCCGGACCACGACATGTAGCTCCGGG 186
QY 510 cgcgtccatgcgaaccgcggagcatgagatgcacagtaaaagacattccagggtggaa 569
Db 187 CGACTCCCTATGACATCCGGAGCCATAGGATCCACAGTAAAGGACATTTCCACGGGTGAA 246
QY 570 tccaagtcataaatagaataaacaagacccttccttaaatctcgaataactgataaacgagc 629
Db 247 TCCAGTCCAAAATGAAAAAACAAGACCTCTCTAAATCTCTGAAAACGTGATAACAGC 306
QY 630 cagataaatccaaatgtagaacacttgcgggaagaagatcttlaacctgttaacctctg 689
Db 307 CAAAAATCCAAATGTAGACACTTTGGGAAAAAGTATTTTACCTGTACTTACTCTG 366
QY 690 tcaacatctgaataacttcaaaagacataaagatcgtgcggagcgagcttgaagat 749
Db 367 TCAACATATCTAAAAAAGCTTCAAAAGACATTTAAGATCTGSGAGGGCGAGTTGAAGAT 426
QY 750 tcttcagcaagaatatcagttactatcttcaataaagaagaagctcaaatctgcacaa 809
Db 427 TTTCAGCAAAATATCAAGTATCTGTTTCAATTAAGAAAGAAAGCAAGTATGCACAA 486
QY 810 ccttggctgcaattctccctgtacacaaagctcgaatctgcataactgcagaagaacctt 869
Db 487 CACGTGGTCAAGTTCCTCTGTACCAAGTCCAGAGTCTGCATATCTGCAGAAACCACT 546
QY 870 caacctaccgcagcatgtagaagtatcatctaaagtcacagacacagtggttaagca 929
Db 547 CACCTCATCCCGCCATGATGGAAGTTCCTTAAGTCTCAGACAGAGTATGTTGGACC 606
QY 930 gaggaatatatgattgtaaaaagctatacaagagacatgatttatctcttcaaatgta 989
Db 607 GAGGAAAGTATTACTGAAAAACGTCGTAAGATCATGATTTATTTATCTCGCTAATATGA 666
QY 990 tatatacaatctctgtcagggagtaaaaactctcatatgataagatataact 1049
Db 667 TATTATCAAAATGCTTAACATGGGAGCTAAAGATCTTCATATGATGACATTAAGTACT 726

QY 1050 acattgaacaaagaagaaagttgttatttactcaagaatcaagtaactcagtaagag 1109
 Db 727 ACATTGAACAGAGAGAGAAAGGTTATGTCTCAAGAAATCAAGTACTTCTGTGAGGG 786
 QY 1110 atggggacaaaagagtggtgtagtgacaaaacaaagaacgaagaagactcaaaaagc 1169
 Db 787 ATTCGGGAAAAAAAGCTGGCACTACAAATACAAAGCGGAGAACGAGGAGCTCAAAAAAGC 846
 QY 1170 ctttgttaaggtggaagatatgagccaactttaggcatttattcttcagctgagca 1229
 Db 847 CTTTTTAAAGGTTAAAGATGTGAACCGGCTTTACAGGCCATTCTACCTTCAGCTGACCA 906
 QY 1230 ataagctttaaataattatcttactcagaagccctgagctccatttgaatgagcaagc 1289
 Db 907 GTGTGCTCTTCTATTAACATATGTCTACGCACAGCCCTTGACGCCATTGTGACATAGAAC 966
 QY 1290 catctagatgcaaaaagaacacacaggtttaaactaagaatcccaacacagatggcgaatg 1349
 Db 967 CATCAGCGGTCCAAAAGAGAGCTCAGCCCTAAACCCAGACCTTAATGCGATGGTGAAT 1026
 QY 1350 atgtggaacctcaattcaactccagttgaaagaagaagaagaagaagatatgtgaat 1409
 Db 1037 GTGTGGGAAACCCCACTCAACTCCAAATTGAAAGAAAGAAAGAAAGGATTTGTGAAT 1086
 QY 1410 gtgtgttcagagaataatgaagttcagaacacacacacacacacacacacacacacac 1469
 Db 1087 GTTGTCTTACAGAAATACGAAAGATCTTGAACCTCCCTTTGTGAGGAAACACAAAAACT 1146
 QY 1470 ttgcacagaagtaacacagatcagttgtgataatgtatcctaagttagttttagt 1529
 Db 1147 TCGCAGACAGATACACAGATTCAGATTTGTGATGATTTGTATCTAAGTATGTTTGGATT 1206
 QY 1530 ttgtggaatatgaagaagac 1589
 Db 1207 TTTGTGAATATGAGAGAGACACACCTTAAAGAAAGAAATTAATACAGCTGTGGCTCT 1266
 QY 1590 ttctcccggttctcgaagtgctcgaagaaagaagacgaagaagaagaagaagaagttgc 1649
 Db 1267 TTTCTTCTGTCTCTCAAAATGCTGAAAAACATCTAACCAAGGAGATCACTCACTGCT 1326
 QY 1650 aacatatcttcagagaagaatgcccaggaagatgatacaacagtggaagagacaga----- 1703
 Db 1327 AGCCCATTTTCCAGAAAGGACATGTTAGAAAGTAAATGGACAACTGCTGAGGAGATTTTC 1386
 QY 1704 -----attccctgataaagaagacccagaaactgaaaaaagctcctgttattt 1754
 Db 1387 AGTGTGAAGATATTCAGAGTGTGAAGATATCCAGAAACCTGAACAAAGGCTGTGTGCT 1446
 QY 1755 cagagcccaatcccccaccccttcaaatgaaatggaagagcttaatgaagaatgaaata 1814
 Db 1447 CAGAACCCATGCTCTACCTTCAACCGGATTTGAAGGAGATGATGAAGAA-----G 1497
 QY 1815 aatgttccatgttaagtaagctgaaagatgacataagaagaagaattttacacagctacac 1874
 Db 1498 CAGCCAGCATGTTAAATGCGCACTGAACCTGACATTAATAAGAGATTTAACAGACTACCTC 1557
 QY 1875 tacataaacaacaagaatgcatcttgcacattccgaacacacacataatgaaatg 1934
 Db 1558 CATGTAAATAATGACAGAGAAATCTGTGATGTTCTCAACATTAATTAATTAATA 1617
 QY 1935 actaga---agaactaaggtgagatcactataaatgttaacatacagcatcgttaca 1991
 Db 1618 GAATATGATTTAAAGAGAGGTTGGTGTGAGTCTGTGAGTGTGACACGGTATGTGAAG 1677
 QY 1992 ttctcgtatttcagatacagataatgagttccaacaaagaagaagatgagatactgagc 2051
 Db 1678 TTTCTCATCTCAGTCCAGAAATAGTCTACCCCAACAAATTAACAGACAGATACAC 1737
 QY 2052 ttctccacaaagaagctcagaagaagaagcccttcattcaatattactgattctg 2111
 Db 1738 ATTTTTCAGCAAAATATTACAAAGAAAGACCTTCATTTTGTGTGTGATGATTCG 1797

QY 2112 gctgataacaataaagcgttcacaagaagacactaactgttcaggcaaaagctccatcc 2171
 Db 1798 ATCTGTGACATTAAATFACATCAAAAAGACAGCTAACAGTGAAGCGGAGACCCATCT 1857
 QY 2172 atactcctcctgaaggaacccaatgaaatgtgacttcaagaatatgataagttactctg 2231
 Db 1858 GTGTCTCTC---AGCAGCCCAATGAGTGTGACACGAGAAATACAGTAACTTCCCTGTG 1914
 QY 2232 gtaaaatctgaagaagaagaataataatagtagaagaatgaagaagaatcctgagac 2291
 Db 1915 GTTAAGATCCAGCGGAAAGTAGAGCTGTATTAGGCGAGAG---AAAAAAATGTGACC 1971
 QY 2292 caaagtcgaatttga---aaaagaactgaaatttatacacagaagaagaagaatc 2348
 Db 1972 CAAGTCTGATTTGATTAATAAAAGAACTGAGTTCTTCTATGTGTGAAGACAGAACTT 2031
 QY 2349 gtagtcaacgggtacagcttcttactagactgttcaagactgtaagaagaaacagaat 2408
 Db 2032 GTGGGTACACAGTACAGTCTCTCTGAGCTTTTCAACAAACGGGAGAAATCAGATT 2091
 QY 2409 ttgttggttccaaagctcacagaagaagaagtgatatagcaatgttttagatattgg 2468
 Db 2092 TTTTGGGTTTACACAGCTACACTGAAGACGTGGCTATGTGATGTGAATGTTGGG 2151
 QY 2469 aagaagaagaatcagaataatcgttcaacagcggttctccgtccctcaactacat 2528
 Db 2152 AGGATTAATAATTCAGTACTCTCTCAACATTTTCTCTTCCCTTCAGCTCTACAT 2211
 QY 2529 ttactgcttcttgaatatttaaaatgcatacttccagaagtgataagatcatct 2588
 Db 2212 TCATTGTGATTTAGTATTAATGTTTTCAGAGTGAACAGTATTTATCTTGAAATTT 2271
 QY 2589 tga-----aatttataataatgataagaaatctttagatatttlltccag 2636
 Db 2272 TATTAATACATATATATTAATCTTAGAGGTCTCTTTTGTGTGTGTTGCTCTTCCAC 2631
 QY 2637 cttgttllacagaccacaatgtaaatatataataataatattgcaatttctacagaat 2696
 Db 2332 CTTTGTGCTAGACACCATGTGTAAATACATAGCA--TGATGTTATTAATTTTTCAGAA 2389
 QY 2697 tgaatcctgttlaaagaagaatcagaataaactgtgactgttcttlltcaataa 2756
 Db 2390 TGAGTACGTCT--ACAAATTAATACAAATTAATCTGTGATTTATTTTACATCAAAAA 2447
 QY 2757 aaaaaaaaaa 2767
 Db 2448 AAAAAAAAAA 2458

RESULT 7
 MIMB4
 LOCUS 2309 bp mRNA ROD 11-NOV-1999
 DEFINITION Mus musculus mRNA for DBF4-related protein, complete CDS.
 ACCESSION AJ003132
 VERSION AJ003132.1 GI:6006562
 KEYWORDS DBF4-related protein.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2309)
 Lepke,M., Putler,V., Stalb,C., Kneissl,M., Berger,C., Hoehn,K.,
 Nanda,I., Schmid,M. and Grunmt,F.
 Identification, characterization and chromosomal localization of
 the cognate human and murine DBF4 genes
 Mol. Gen. Genet. 262 (2), 220-229 (1999)
 JOURNAL MEDLINE 99444905
 MEDLINE 2 (bases 1 to 2309)
 REFERENCE
 AUTHORS Lepke,M.
 TITLE Direct Submission
 JOURNAL Submitted (17-DEC-1997) Institute of Biochemistry, Am Hubland,
 Wuerzburg D-97074, Germany
 FEATURES Location/Qualifiers

repeat_region 5950..6035 /rpt_family="MERL_type"
 repeat_region 6360..6642 /rpt_family="Alu"
 repeat_region 7010..7308 /rpt_family="Alu"
 repeat_region 9806..10112 /rpt_family="Alu"
 repeat_region 10303..10498 /rpt_family="Alu"
 repeat_region 10507..10807 /rpt_family="L1"
 repeat_region 10861..10978 /rpt_family="Alu"
 repeat_region 11533..11836 /rpt_family="L1"
 repeat_region 11959..12079 /rpt_family="Alu"
 repeat_region 13992..14290 /rpt_family="Alu"
 repeat_region 15571..16139 /rpt_family="L1"
 repeat_region 16142..16438 /rpt_family="Alu"
 repeat_region 16500..16779 /rpt_family="Alu"
 repeat_region 16933..17077 /rpt_family="Alu"
 repeat_region 17850..18392 /rpt_family="L1"
 repeat_region 18480..18769 /rpt_family="Alu"
 repeat_region 18908..19366 /rpt_family="L1"
 repeat_region 21782..21896 /rpt_family="L2"
 repeat_region 22048..22156 /rpt_family="L2"
 repeat_region 24873..25250 /rpt_family="MERL_type"
 repeat_region 26032..26309 /rpt_family="MauR"
 repeat_region 26342..26817 /rpt_family="L2"
 repeat_region 27147..27355 /rpt_family="MauR"
 repeat_region 27381..29443 /rpt_family="Alu"
 repeat_region 29444..29561 /rpt_family="L1"
 repeat_region 31770..31809 /rpt_family="Alu"
 repeat_region 32332..32720 /rpt_family="MIR"
 repeat_region 32731..32953 /rpt_family="MauR"
 repeat_region 33094..33327 /rpt_family="Alu"
 repeat_region 33643..33941 /rpt_family="MERL_type"
 repeat_region 34000..34088 /rpt_family="Alu"
 repeat_region 34233..34602 /rpt_family="MauR"
 repeat_region 38520..38721 /rpt_family="MIR"
 repeat_region 38730..38849 /rpt_family="Alu"
 repeat_region 40214..40605 /rpt_family="L2"
 repeat_region 41565..41638 /rpt_family="L1"
 repeat_region 42099..42347

repeat_region 42470..42784 /rpt_family="MIR"
 repeat_region 42787..43087 /rpt_family="L1"
 repeat_region 43588..43892 /rpt_family="Alu"
 repeat_region 44036..44116 /rpt_family="Alu"
 repeat_region 44562..45028 /rpt_family="L1"
 repeat_region 45191..51321 /rpt_family="L1"
 repeat_region 52546..52617 /rpt_family="L2"
 repeat_region 53530..53581 /rpt_family="L2"
 repeat_region 54254..54761 /rpt_family="L2"
 repeat_region 54759..55585 /rpt_family="L1"
 repeat_region 55659..55657 /rpt_family="L1"
 repeat_region 56134..58298 /rpt_family="Alu"
 repeat_region 59008..59181 /rpt_family="MIR"

Query Match 43.0% Score 1194.2; DB 9; Length 123331;
 Best Local Similarity 99.7% Pred. No. 1.1e-216;
 Matches 1196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1563 aagaataaataacagtggtgagatcccttcctctggttcctgcaagtgctcgaanaaga 1622
 1623 ctgaacaaagaagaagtggaatcgacacatattctcagaanaagttccggaagatg 1662
 94286 AAGAAATGAAATACAGTGTGATCCCTTCTCCTGTTCTGCAAGTGTCTGAAAAGA 94227
 94226 CTGAACAAAGCAAAAGTGGAATTCGAACATATTCTCGAAAAGTTCCCGAAGATG 94167
 1683 atacaacagtaagaagacagaattctcgtataagaagaccggaactgaanaaagc 1742
 94166 ATACACAGCTAAGACAGCAAAATTCCTGTATTAAGAGACCAGAACTGAAAAAGC 94107
 1743 tcccttattatcgagagccatcccccactcaaatgaattgagaggttaagaga 1802
 94106 TCCCTATTATTTCAGACCCATCCGCCCTTAAATGAATGAGGGCTTAATGAGA 94047
 1803 aaatagtaataaattgttcattttaagtaagcctgaagatgacataagaacgaattta 1862
 94046 AATATAGATTAATATGTTCCATGTTAAGTACGCTGAAGATGACATTAAGCAGAAATT 93867
 1863 cccagctacccctacataaacaacaggaatgcatcttgaacatttcggaacacacat 1922
 93986 CACACCTACCTCTACATPAACAAACAGAAATGATCTTGACATTTCCGAAACACAT 93927
 1923 taagtgaataaagtaagtaagaactaaggtagatcactataaattgaatacaggaat 1982
 93926 TAAGTGAATAAGACTTGAAGAACTAAGGTGATCACTATTAATTAATGATACAGGCA 93867
 1983 cgtacagcttcctgatttcagtaacagataaagtagtgaattccaacaaacagaagtcag 2042
 93866 CGTACAGCTTCTGATTTTCAGTACAGATTAATGATGATCTCAACCAACCAAGTACG 93807
 2043 atacagcttccttcacagcaagaatcctcaagaagaagaacctcattcaatattac 2102
 93806 ATACAGCTGCTTTTCAGCAAGATCTCAAGAAAGACCTTCATTAATATTATTCAC 93747
 2103 atgatttcgctcgtataaataaataacagttcaacagcaccactcgttcaggaagag 2162
 93746 ATGATTTCGCTCTGTAACAATAAACGTTCAAGAGCACCTAAGTTCAGGCAAGAG 93687
 2163 ctcacatcaactccctccggaggaaacccaatgaattgacttaagaataatgagtagt 2222

```
Db 93686 CTCATTCACATCTCTCTCTGAGAACCCAAATGAATGTGACTTCAAGAAATATGATGATTT 93627
QY 2223 taccttcggtgtaaaatacatggaagaatggaataatataattaggaagaaatagaagaaga 2282
Db 93626 TACCTTCGTGTAATAATACATGAAAAAGTAAATATATTAGACGAAATATAAAAAGAAA 93567
QY 2283 atctggaacccaatgctgaattgataaagaactgaattttcacagaagaagaaca 2342
Db 93566 ATCTGGAACCAAAATCTGATTTGATTTAAAGACATGATTTTTCACAGAGAGAAAAACA 93507
QY 2343 gaattgtagtccacggtacacgtcttactagactgtttcagaactgtagaagaagaat 2402
Db 93506 GAATTTGTAGTTACCGGTACAGCTTCTTACTAGACTGTTTCAGACTAGTGAAGAAAT 93447
QY 2403 cagaatttttgggtttcacaagctcacagaagaagaatggtatagcaatgttttagata 2462
Db 93446 CAGAAATTTTGGGTTTCACAAACCTACACAGAAAAGAGGTAAATGCAATGTTTAAAGATA 93387
QY 2463 ttgggaagagaaataatcagaataatctgttacaagcgltttccgtccctccaact 2522
Db 93386 TTTGGGAAGAGAAAAATTCAGATATCTGTATACAGCGTTTTCGTCCCTTCAACTT 93327
QY 2523 ctacattactggtctttagaatttaaaaaatgcatactttcagaagtgtataagataca 2582
Db 93326 CTACATTTACTGGCTTTTAGAATTTAAAAAATGCATACTTTTCAGAGTGAATAGATCA 93267
QY 2583 tattctgaaatttttataataatgtaagaaatccttaaggaatttttttccaagcttgt 2642
Db 93266 TATTTCTTAAAAATTTTAAATAATGTATGGAATTCCTTAGGATTTTTCACAGCTTTGT 93207
QY 2643 ttacagaccacaatgtataataataataaataatttgaacttttctagaagaattgaata 2702
Db 93206 TTACGACACCCAAATTAATATTTAAAAAATTAATTTTGAATTTTCTACAGATTTGAATA 93147
QY 2703 cctgttaagaanaaattcagaataaactgtgactgtgctgttttactaataaaaaa 2761
Db 93146 CCTGTTAAAGAAAATTAACAGATTAACCTGTGACTGCTGTTTACATTAATATATA 93088

RESULT 9
AC024947 163031 bp DNA PRI 24-AUG-2000
LOCUS Homo sapiens chromosome 01 clone RP11-200110, complete sequence.
DEFINITION AC024947
ACCESSION AC024947.4 GI:9887599
VERSION HTG.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 163031)
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163031)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 3 (bases 1 to 163031)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On Aug 24, 2000 this sequence version replaced gi:8569066.
FEATURES
source location/Qualifiers
1..163031
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="01"
/clone="RP11-200110"
/clone_lib="RPCI-11"
```

```
BASE COUNT 48079 a 30305 c 32059 g 52588 t
ORIGIN
Query Match 43.0%; Score 1194.2; DB 9; Length 163031;
Best Local Similarity 99.7%; Pred. No. 1.1e-216;
Matches 1196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1563 aagaataaataacagtggttgatcccttccctgttctcgaagtgtccggaagaaga 1622
Db 35370 AAGATATATAAATACATGTTGATCCCTTCTCTTCTTCGAAATGTCGAAAAACA 35429
QY 1623 ctgaacaaaggaaagaagtggaatgcaacataattcagaagaatgccagaagatg 1682
Db 35430 CTGAACAAAAGAAAAAGTGGAAATTCACATATTTCTCGAAGAAGTTGCCAGAAAGATG 35489
QY 1683 atacacagtggaagagcagaatttccgtlataaagagacccaggaactgaaaaaagc 1742
Db 35490 ATACAACTGAAGAGAGAGCAATTTCTGTATAAAGAGACCCAGAAACTGAAAAAAGC 35549
QY 1743 tccgttttattcagaagcccatccccccttcaaatgaatgtagaggtctaatgaga 1802
Db 35550 TCCGTGTTATTTTCAGAGCCCATCCCCACCTTCAATATGATGAGAGGCTTAAATGACA 35609
QY 1803 aatgagtaataatgttccatgttaagtagcagctgaagaatgacataagaagaattta 1862
Db 35610 AATAGTATATAATGTTCCATGTTTAACTAGTACAGCTGAAGTGCATTAAGACAGATTTTA 35669
QY 1863 cacagtaaccttacaataaacaacaggaatgcatcttctgacattccgaacacacat 1922
Db 35670 CACAGTACTCTTACATTAATAAACAACAGAAATGCAATTTCTTGACATTTCCGAACACACAT 35729
QY 1923 taagtaaaatagaacttaagaagaactaaggtagatcactataaataatgatacagaagcat 1982
Db 35730 TAAGTAAAAATGACTTGAAGAAGAACTAAGGATGATCATATAATGTATACATACAGGCAAT 35789
QY 1983 cgtacagtttcttgatcttcagtaacagataatagtgatcctcaacaaacagaagtcag 2042
Db 35790 CTGTACATGTTTCTGTGATTTCAGTACAGATATAATGTGATCTCACCAAAAACAGAAATCAG 35849
QY 2043 atactgtcttltccagaagaagatcacaaggaagaagaccttcaataatttactc 2102
Db 35850 ATACGTGCTGTTTTCACAGAAAGATCTCAGAGAAAAGACCTTCAATCAATATTATTCAC 35909
QY 2103 atgattcgtgtctgataacaataaacaagttcacagaagaacacctgaactgttcagcaag 2162
Db 35910 ATGATTCGTGCTGATTAACAAATTAACAGTTCACAAAGCACCTTAACGTTCACGCAAAAG 35969
QY 2163 ctccattcattactcctctctgaggaacccaatgaatgtgacttcaagaataatgataatc 2222
Db 35970 CTGCATTCATACTCTCTCTGAGGAACCCAAATGAATGTGACTTCAAGAAATATGATAGTT 36029
QY 2223 taccttcggttaaaatacatcgaaagtgaaaataatattaggaagaaatagaagaaga 2282
Db 36030 TACCTTCGTGTAATAATACATGAAAAAGTAAATATATTAGACGAAATATGAAGAAACAA 36089
QY 2283 atctggaacccaatgctgtaattgataaagaactgtaatttattacaagaagaagaaca 2342
Db 36090 ATCTGGAACCAAAATGCTGAATTTGAATTAAGAACTGAATTTATTAACAAAGAGAAACAC 36149
QY 2343 gaattgtagtccacggtacacgtcttactagactgtttcagaactgtagaagaagaat 2402
Db 36150 GAATTTGTAGTTACCGGTACAGCTTCTTACTAGACTGTTTTCAGACTAGTGAAGAAAT 36209
QY 2403 cagaatttttgggtttcacaagctcacagaagaagaatggtatagcaatgttttagata 2462
Db 36210 CAGAAATTTTGGGTTTCACAACTACACAGAAAAGAGGTATATATCAATGTTTAGATA 36269
QY 2463 ttgggaagagagaataatcagataatctgttaacagcgltttcctgtccctccaact 2522
Db 36270 TTTGGGAAGAGGAAAATTAACAGATTAATCTGTAAACACGTTTCTGTGCTCCCTTCAACTT 36329
QY 2523 ctacattactggtctttagaatttaaaaaatgcatactttcagaagtataagataca 2582
```


|||||
Db 36330 CTACATTTTACGCTTTTACGATTTTAAATTAATGATCTTTTACAGAGTATDAGATCA 36389
2583 taattctgaatttttaataatgtaagaaattcttaagattttttacagacttctg 2642
Db 36390 TATCTTGAATTTTATTAATATGATGGAATTTCTTAAGATTTTTCACGCTTGT 36449
2643 ttacagaccacaatgtaataatataataatatttgcattttctacagaattgaa 2702
Db 36450 TTACAGACCCCAATGTAATTAATTAATTAATTTTCTACAGATTCATAA 36509
QY 2703 cctgtttaaagaataattacagaaataactgtgactgtgcttcttacaataaaaa 2761
Db 36510 CCTGTTAAAGAAAATTAACAGATTAATTAATTTGACTGCTGTTTACATTAATATA 36568

RESULT 10
AC004958 163889 bp DNA HTG 12-JUN-1998
LOCUS Homo sapiens clone D01096F22, *** SEQUENCING IN PROGRESS ***, 6
DEFINITION unordered pieces.
AC004958
VERSION AC004958.1 GI:3213039
KEYWORDS HTG; HTGS_PHASEL.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 163889)
Waterston, R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 163889)
Waterston, R.H.
Direct Submission
Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108 USA

COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1462: contig of 1462 bp in length
* 1463 1480: gap of unknown length
* 1481 3064: contig of 1584 bp in length
* 3065 3082: gap of unknown length
* 3083 4710: contig of 1628 bp in length
* 4711 4728: gap of unknown length
* 4729 6530: contig of 1802 bp in length
* 6531 6547: gap of unknown length
* 6548 16429: contig of 9882 bp in length
* 16430 16446: gap of unknown length
* 16447 163889: contig of 147443 bp in length.

FEATURES
SOURCE
1..163889
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="D01096F22"

BASE COUNT 49499 a 30566 c 31570 g 52166 t 88 others
ORIGIN

Query Match 43.0%. Score 1194.2. DB 2. Length 163889;
Best Local Similarity 99.7%. Pred. NO. 1.1e-216.
Matches 1196; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1563 aaagaataaatacagatgttgatcccttctcctgtttctcgaagtctctgaaaga 1622
|||||
Db 74946 AAAGAAATTAATACAGTGTGATCCCTTCTCCTCTTCGAGTGCTCGAAGAAAGA 75005

QY 1623 ctgaagaaaggaagatggaattgtcacatacttcagaagaatgtccagaagatg 1682
Db 75006 CTGACCAAAAGGAAAAAGTGGAAATTCACATATTTCTCAGAAATTTGCCAGGAATG 75065
QY 1683 atcaacaattgagaagagaagaattcctgtataagaagaccagaagaactgaagaaagc 1742
Db 75066 ATCAACAGTGAAGAGAGCAAGATTTCTGTATTAAGAGACCAGGAATCTGAAAAAGCC 75125
QY 1743 tctgttattcagaagcccatcccccacccctcaatgaattgagaaggttaatgaag 1802
Db 75126 TCTGTTATTTTCAGAGCCATCCCCACCTTCATTAATTAATGAGAGGCTTAATGAGA 75185
QY 1803 aaatgaaataaatgttccatglttaagtaacagctgaagatgaataagaagaattta 1862
Db 75186 AAATGACTAATTAATGTTCCATGTTAAGTACAGCTGAAGATGACATTAAGACAGAAATTTA 75245
QY 1863 caccgtaccctctacataaaaaacaagaaatgtcattcttgacattccagacaacat 1922
Db 75246 CACAGTACTCTTACATATAAAACAACAGAAATGCAATTCGTGACATTTCCGAAACACAT 75305
QY 1923 taagtgaataatgacttagagaagactaaggatgaactataaattgatacacaaggcat 1982
Db 75306 TAACTGAATGACTTAGAAGAACTAAGCATAGATCACTATAATGTAACATACAGCAT 75365
QY 1983 ctgtaaagtgttctgtatctcagtagaagaataagtgatctcagaacaagaagatcaag 2042
Db 75366 CTGTACATGTTTCTGATTTCACTACAGATTAATGATGATCTCACCAAAACAGAACTCAG 75425
QY 2043 atactgtctttctcagaagaagatctcagaagaagaagaccttctcaatattactc 2102
Db 75426 ATACGTGCTTTTTCAGAGAAAGATTCAGAGAAAGACCTTCATTAAATTTACTC 75485
QY 2103 atgattctgtctgtataacaataaacagttcaagaagaccataatgttcaagaagaag 2162
Db 75486 ATGATTTCTGTCGTATTAATAAATTAACACTTCAAGAGCACTTACTGTTCCAGCAAAAG 75545
QY 2163 ctcaatctcaactcctctctgagaagaccgaatgaatgtgaactcaagaatagtatgtc 2222
Db 75546 CTCCATTTCCATCTCTCTCTGAGAAACCAATGAATGTGACTTCAAGAAATATGATATCT 75605
QY 2223 tacctctgtgtaataacacacgcgaagaatgaataataatataagacgaatagaagaagaa 2282
Db 75606 TACCTTCGTGTAATAATACATCGAAAGTGAATTAATTAAGACCAATTAAGAAAGAA 75665
QY 2283 atctggagaaccaatgtcgaatttgaataaagaactgaattatcacagaagaagaaca 2342
Db 75666 ATCTGGAACCAATGCTGAATTTGATTAAGACAGTAATTTTACACAAGAAAGAAACA 75725
QY 2343 gaattgtaattcaacggatgaagttcttactagaactgtttcagaactagtagaagaagaa 2402
Db 75726 GAATTTGTATTCACCGGTACAGTCTTTACTAGACTGTTTTCAGACTAGTGAAGAAAT 75785
QY 2403 cagaatttttgggtttcaacagcttacaagaagaagaagtgatatacaatgttttaagta 2462
Db 75786 CAGAAATTTTGGGTTTACAGCTACAGAAAGAGTGTATATGCAATGTTTAAGATA 75845
QY 2463 ttctggagaagaaatcaatcaatcgttacaacaggttttctcgtcccttaact 2522
Db 75846 TTTGGAGAGAGAAATTCAGATTAATCTGTTTACAGCTTTTCTCCTCCCTTCAACTT 75905
QY 2523 ctacattactgtgcttttgaatttaaaaaatgacattcttcaagaagatgaatgaatca 2582
Db 75906 CTACATTTTACTGCTTTTGAATTAATAAATTAATTAATTTTCAAGAAATGAATGAATCA 75965
QY 2583 taattctgaatttttaataatgtaagaaattcctaagatttttttaccagcttctg 2642
Db 75966 TATCTTGAATTTTATTAATATGATGCAAAATTTAGATTTTATTAACAGCTTGT 76025
QY 2643 ttacagaccacaatgtaataatataataatatttccaatttctcagaatgtgaata 2702
Db 76026 TTACAGACCAAAATGTAATTAATTAATAATTAATTAATTTTCAATTTTCTACAGAAATGAATA 76085

QY 2703 cctgtttaaagaataatcacagaataactgtgactgctgtgttcttaacataaaaaa 2761
|||||
Db 76086 CCTGTTAAAGAAAAATTAACGAATTAACCTTGACGTGCTGCTTCTTTACATTATATATA 76144

RESULT 11
AK022969 2276 bp mRNA PRI 29-SEP-2000
LOCUS Homo sapiens CDNA FLJ12907 fis, clone NT2RP2004396, highly similar
DEFINITION to Homo sapiens mRNA for activator of S phase Kinase.
ACCESSION AK022969
VERSION AK022969.1 GI:10434668
KEYWORDS oligo cloning; fis (full insert sequence).
SOURCE Homo sapiens teratocarcinoma cell line:NT2 CDNA to mRNA,
clone_11b:NT2RP2 clone:NT2RP2004396.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y.,
Matsunawa,H., Ishii,S., Kawai,Y., Salto,K., Yamamoto,U.,
Wakamatsu,A., Makamura,Y., Nagahari,K., Masuko,Y. and Sasaki,N.
TITLE NEDD human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 2276)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao
Isogai, Helix Research Institute, Genomics Laboratory: 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp,
Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDD human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; CDNA full insert
sequencing; Research Association for Biotechnology; CDNA library
construction, 5'- 5' and one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Source
Location/Qualifiers
1..2276
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone="NT2RP2004396"
/clone_11b="NT2RP2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."

BASE COUNT 793 a 381 c 421 g 681 t
ORIGIN

Query Match 39.6%; Score 1101.6; DB 9; Length 2276;
Best Local Similarity 99.6%; Pred No. 3.5e-199;
Matches 1104; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1654 tatttctcagaagaatctgcaggaagaatgatacaacagtgtaagagagaagaattcctgta 1713
|||||
Db 1 TATTTCTCAGAAAGATTGCCAGAGATGATACACAGTAAAGAGAGAGATTTCCGCTGA 60

QY 1714 taagaagaccaggaactgtaaaaaagctcgttatttccagagccatcccccacc 1773
|||||
Db 61 TAAAGGACCCAGGAAATGAAAAAGCTCCGTTTATTTCCAGAGCCATCCCCACC 120

QY 1774 ttcaaatgattgagaaggcttaataagaataatgagtaataaaatttccatgttaagtac 1833
|||||
Db 121 TTCAAATGATTTGAGGGCTTATATGAGAAATGAGTAATTAATGTTCCATGTTAAATAC 180

QY 1834 agctgaagatgacataagaacagatttccacagctaccctctacataaaaaaacagga 1893
|||||
Db 181 AGCTGAAGATGACATAGACAGAAATTTTCACAGCTACCTCTACATATAAAACAAACAGGA 240

QY 1894 atgcattcttgacatttcggaacacacataaagtgaanaatgacttagaagaactaaggt 1953
|||||
Db 241 ATGCATTTCTTGACATTTCCACACACACATTAAGTGAAGAAATGACTTGAACACACTAAAGGT 300

QY 1954 agatcacataaatgttaacatacagagcatcgtacatgcttccatctcgtatccagataa 2013
|||||
Db 301 AGATCACTATAAATGTATACATACAGAGCATCTGTACATGTTTCATGATTTCCATACAGATTA 360

QY 2014 tagtggatctcaacaaacagaagtcagatactgcttctccagaaagatcaca 2073
|||||
Db 361 TAGTGATCTCAACCAAAACAGAAATGAGTACTGTCTTTTCCAGCAAAAGATCTCAA 420

QY 2074 ggaagaagaccttcattcaatatttactatgatttctgctgctgataacaaataaagcttc 2133
|||||
Db 421 GGAAGAGACCTTCATTCAATATTACTCATGATTCGGTGTGATTAACAATTAACACTTC 480

QY 2134 acaagagcacctactgttcaagcaaggctccatccatccatccctcctgaagaaaccaa 2193
|||||
Db 481 ACAAGAGCACCTACTACTGTTCAGGCAAAAGCTCCATTCATACCTCTCTGAGAACCCAA 540

QY 2194 tgaatgtgactccaagaataatgatatagtttaaccttcgtgtaaaatatacgcgaagtga 2253
|||||
Db 541 TGAATGTGACTTCAAGATATGATAGTATACCTCTGTGTAATAATCATCGAAAGTGAA 600

QY 2254 aataaataatagagcgaataatgaagaanaatctggaaccaaagtctgaattatgataaag 2313
|||||
Db 601 AATTAATATTGAGACGAATATGAAAAAGAAATCTGGAACCAATGCTGAATTTGATTAAG 660

QY 2314 aactgaaattatcacacaagaagaacagaattgttagtccaacggttacaagcttact 2373
|||||
Db 661 AACTGAATTTATTAACACAGAGAAAAACAGATTTGATGTTCAACGGTTACAGTCTTACT 720

QY 2374 agactgttccagactgtaagaagaatcaagaatttgggttccaaagctcacaga 2433
|||||
Db 721 AGACTGTTTCAGACTGTGTGAGAGAAATGAAATTTTGGGTTTCAACAGCTCACAGAGA 780

QY 2434 aaagagagatataatgacatttttagatatttggaagagaaatccagataatcgt 2493
|||||
Db 781 AAAGAGGTATATGACAAATGTTTATTTGGGAAAGAGAAATTCAGATTAATCTGTT 840

QY 2494 aacagcgttttctcgcgcccttcaacttccatctacattacgcttcttagaatttaaaaa 2553
|||||
Db 841 AACAGCGTTTTTCTCGGCCCTTCACTTCTACATTTACGCGCTTTTGAATTTAAAAAA 900

QY 2554 tgcatacttttcagaaagtgataagagatcatalctctgaaatttttataaataatgata 2613
|||||
Db 901 TGCATACTTTTCAGAAAGTAAAGATCATATCTTGAATTTTATTAATATGATGGA 960

QY 2614 aattcttaggaattttttaccagcgttggtttccagaccaccaatgtaatatataaataa 2673
|||||
Db 961 AATTCCTTAGGATTTTATTTTACCAGCTTGTGTTTACAGACCCAAATGTAATTTAAAAATTA 1020

QY 2674 atatttgcatttctcagaatgtaatacctgtttaagaanaaatcagaataactgt 2733
|||||
Db 1021 ATATTTCATTTTCTACAGAAATTTGAATTTACCTGTTAAAGAAAAATTTACAGAAATAACTTG 1080

QY 2734 tgaactgctctgttttatacaataaaaaa 2761
|||||
Db 1081 TGACTGCTCTGTTTATCATTTATATATA 1108

RESULT 12
AC091341/c
LOCUS AC091341 232903 bp DNA HTG 18-APR-2001
DEFINITION Rattus norvegicus clone CH230-1C19, WORKING DRAFT SEQUENCE, 35
unordered pieces.
ACCESSION AC091341
VERSION AC091341.1 GI:13661907
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE AUTHORS

TITLE	JOURNAL
REFERENCE	AUTHORS
TITLE	JOURNAL
COMMENT	
<hr/>	
<p> Aldredicks, S.L., Amaratunge,H.C., Ate,U.R., Banks,T., Barbatta,J., Benito, J., Bimac, K., Blankenburg, R., Bonnin,D., Bouck, J., Bowles, S., Brileva,M., Brown,E., Brown,M., Bryant,C.N., Buhey,C., Burich,P., Butrell,C., Butrell,K.L., Byrd,N.C., Carson,T.P., Fisher, Carter,M., Cavaosz,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M., Carroll,D., Dachorne,S.R., David,R., Davila,M.L., Davis,C., Davy,L., Dederich,D.A., Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H., Douhanalt,C.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Fretutz,P., Gabriel,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Hollaway,C., Hollands,B., Homsfi,F., Howard,S., Huber,J., Hulyk,S., Hunne,T.S., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolyvet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kurushi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtenberg-O., Iley,C., Liu,J., Liu,W., Lounsgied,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Mcheshwari,M., Mapua,P., Martin,R., Marindale,A., Matinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Meig,G., Metker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newkison,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenko,S., Ogulu,M., Okunonu,G., Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Peter,L., Peters,L., Pichens,R., Plims,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojnokken,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S., Scott,G., Shen,H., Shoshtari,N., Slason,I., Sodogren,E.,Sonalike T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabori,P., Tamelisso,A., Tametisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,I., Terrell,B., Thomas,R., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalona,D., Vinson,R., Wall,R., Wang,S., Ward-Moorez,S., Warren,R., Washington,C., Washington,S., Williams,G., Williamson,A., Wleczka,R., Wooden,S., Wolley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zortilla,S., Neilson,D. and Gibbs,R. </p>	
Direct Submission	
Unpublished	
2 (bases 1 to 232903)	
Worley,K.C.	
Direct Submission	
Submitted (18-APR-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	
----- Genome Center -----	

TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 232903)
AUTHORS	Motley, K. C.
JOURNAL	Direct Submission
COMMENT	Submitted (18-PPR-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
	----- Genome Center

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)

NOTE: This is a 'working draft' sequence. It currently consists of 35 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	28075: contig of 28075 bp in length
28076	28076: gap of unknown bp in length
28176	42563: contig of 1438 bp in length
42564	42563: gap of unknown length
42654	62492: contig of 19769 bp in length
62493	62492: gap of unknown length
62533	79400: contig of 16868 bp in length
79401	79500: gap of unknown length
79501	92855: contig of 13335 bp in length
92856	92855: gap of unknown length
92857	103757: contig of 10802 bp in length
103758	103857: gap of unknown length
103858	117059: contig of 13202 bp in length
117059	117159: gap of unknown length
117160	126002: contig of 8843 bp in length
126003	126102: gap of unknown length
126103	134408: contig of 8306 bp in length
134409	134508: gap of unknown length
134509	144061: contig of 9553 bp in length
144062	144161: gap of unknown length
144162	153305: contig of 9144 bp in length
153306	153405: gap of unknown length
153406	161399: contig of 7994 bp in length
161400	161499: gap of unknown length
161500	168248: contig of 6779 bp in length
168249	168378: gap of unknown length
168379	175475: contig of 7097 bp in length
175476	175575: gap of unknown length
175576	180962: contig of 5377 bp in length
180963	181052: gap of unknown length
181053	185776: contig of 4724 bp in length
185777	185876: gap of unknown length
185877	189454: contig of 3578 bp in length
189455	189554: gap of unknown length
189555	193453: contig of 3951 bp in length
193456	193585: gap of unknown length
193586	198288: contig of 4713 bp in length
198289	198388: gap of unknown length
198390	201470: contig of 3072 bp in length
201471	201570: gap of unknown length
201571	205248: contig of 3728 bp in length
205249	205388: gap of unknown length
205390	208663: contig of 3185 bp in length
208664	208663: gap of unknown length
208684	211371: contig of 2688 bp in length
211372	211471: gap of unknown length
211472	215483: contig of 4012 bp in length
215484	215583: gap of unknown length
215584	217293: contig of 1710 bp in length
217294	217393: gap of unknown length
217394	218894: contig of 1501 bp in length
218895	218994: gap of unknown length
218995	220700: contig of 1796 bp in length
220701	220800: gap of unknown length
220801	223253: contig of 2263 bp in length
223254	223354: gap of unknown length
223355	224310: contig of 1257 bp in length
224311	224610: gap of unknown length
224611	225939: contig of 1349 bp in length
225940	226059: gap of unknown length
226060	227888: contig of 1823 bp in length
227889	227988: gap of unknown length
227989	229034: contig of 1046 bp in length
229035	229134: gap of unknown length
229135	230482: contig of 1348 bp in length


```

repeat_region 5768. .7030
/rept_family="MaLR"
repeat_region 7272. .7610
/rept_family="Retroviral"
repeat_region 7403. .7628
/rept_family="MER4-group"
repeat_region 7650. .7696
/rept_family="Retroviral"
repeat_region 8013. .8755
/rept_family="L1"
repeat_region 8777. .9094
/rept_family="Alu"
repeat_region 9389. .10058
/rept_family="MER2L-g"
repeat_region 10069. .10231
/rept_family="L1"
repeat_region 10626. .11262
/rept_family="L1"
repeat_region 11263. .11827
/rept_family="Retroviral"
repeat_region 11828. .12240
/rept_family="L1"
repeat_region 12241. .12532
/rept_family="Alu"
repeat_region 12533. .12548
/rept_family="L1"
repeat_region 12549. .12849
/rept_family="Alu"
repeat_region 12850. .13169
/rept_family="L1"
repeat_region 13205. .13372
/rept_family="Alu"
repeat_region 13376. .13881
/rept_family="L1"
repeat_region 13882. .14120
/rept_family="L1"
repeat_region 14084. .14283
/rept_family="L1"
repeat_region 14286. .14760
/rept_family="L1"
repeat_region 14765. .15133
/rept_family="MaLR"
repeat_region 15316. .15669
/rept_family="MaLR"
repeat_region 17202. .17287
/rept_family="L1"
repeat_region 17818. .18025
/rept_family="L1"
repeat_region 18030. .18105
/rept_family="L1"
repeat_region 18115. .18424
/rept_family="Alu"
repeat_region 18791. .18835
/rept_family="(CAT)n"
repeat_region 18888. .19165
/rept_family="Alu"
repeat_region 19778. .20821
/rept_family="Retroviral"
repeat_region 20875. .20991
/rept_family="MIR"
gene
<21983. .42089
/gene="WUGSC:H_RG060N22.1"
join("<19983. .22046,24028. .24134,28390. .28457,30361. .30485,
31216. .31389,34475. .34584,36624. .36705,41528. .41608,
41977. .42089)
/gene="WUGSC:H_RG060N22.1"
/note="mitochondrial carrier protein-like; similar to
Q09461 (PDB:q2497990); H_RG060N22.1"
/codon_start=1
/evidence=not experimental
/protein_id="AAC82534.1"
/db_xref="GI:2588623"
/translation="IYTPIDVVKIRIQANNNPLPKGCFVYSNGIMDLCYCEEGGNK

```

[illegible]

AUTHORS Williams, L.T., Escobedo, J., Jniss, M.A., Garcia, P.D., Klinger, J., Kasam, A., Reinhard, C., Randazzo, F., Kennedy, S.C., Pot, D., Lamson, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M., Labat, I., Leshkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.

TITLE Human genes and gene expression products

JOURNAL Patent: WO 0102568-A 639 11-JAN-2001; CHIRON CORPORATION (US); HYSEQ, INC. (US)

FEATURES

source 1. 685 /organism="Homo sapiens" /db_xref="taxon:9606"

misc_feature 1. 685 /note="n = A,T,C or G"

BASE COUNT 238 a 127 c 125 g 187 t 8 others

ORIGIN

Query Match 18.5% Score 513.8; DB:6; Length 685; Best Local Similarity 94.9%; Pred. No. 1,2e-87; Matches 560; Conservative 0; Mismatches 24; Indels 6; Gaps 3;

QY 558 tccagggtggaatcccaagtcacaaatgaaaaaacagaccatcttgaatctctgaaaa 617
DB 1 tccagggtggaatcccaagtcacaaatgaaaaaacagaccatcttgaatctctgaaaa 60

QY 618 ctgatacagcgacagaaatcccaatgtaagccacttggggaagaatttactcttg 677
DB 61 ctgatacagcgacagaaatcccaatgtaagccacttggggaagaatttactcttg 120

QY 678 actacactctctccatctctgaaaaactcaaaagacataagagatctggagagc 737
DB 121 actacactctctccatctctgaaaaactcaaaagacataagagatctggagagc 180

QY 738 gaattgaaatctctcagcaagataatcagttatcttatttcaataagaagaacta 797
DB 181 gaattgaaatctctcagcaagataatcagttatcttatttcaataagaagaacta 240

QY 798 aatttgacaaactctggtgacatctctctgtaagtcacagatcagatctcacta 857
DB 241 aatttgacaaactctggtgacatctctctgtaagtcacagatcagatctcacta 300

QY 858 cagaacaccatctcactcagcagcagatgagagatcatttaagtcacagacag 917
DB 301 cagaacaccatctcactcagcagcagatgagagatcatttaagtcacagacag 360

QY 918 tgggttaagcagaggaatattatggttgaaagaactcaagagacagatcttctc 977
DB 361 tgggttaagcagaggaatattatggttgaaagaactcaagagacagatcttctc 420

QY 978 ctcaaatagatataatataatgacatctgtaagtgagagtaaaatctctcaatg 1037
DB 421 ctcaaatagatataatataatgacatctgtaagtgagagtaaaatctctcaatg 480

QY 1038 acattgatactcattcagacaaagaagaagaagtgatattactcaagaatacagta 1097
DB 481 acattgatactcattcagacaaagaagaagaagtgatattactcaagaatacagta 536

QY 1098 cttaagaagaagatggggcagaagaagtgatggatggatggatggatggatggat 1147
DB 537 cttaagaagaagatggggcagaagaagtgatggatggatggatggatggatggat 584

RESULT 15

AC074175 177899 bp DNA HTG 15-JUL-2000

LOCUS Mus musculus chromosome 3, clone RP23-348B11 strain C57BL/6J, WORKING DRAFT SEQUENCE, 15 unordered pieces.

AC074175 AC074175.1 GI:9211286

VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.mouse@nih.gov

Center project name: vc

Center clone name: 348B11

----- Summary statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 171084 bases at least Q40

Consensus quality: 172951 bases at least Q30

Consensus quality: 174122 bases at least Q20

Insert size: 187000; agarose-fp

Insert size: 202000; pulse-field-gel

Insert size: 176499; sum-of-contigs

Quality coverage: 4.87x in Q20 bases; agarose-fp

Quality coverage: 4.50x in Q20 bases; pulse-field-gel

Quality coverage: 5.16x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 2463: contig of 2463 bp in length

* 2464: 2463: gap of unknown length

* 2465: 2464: contig of 2204 bp in length

* 4867: 2465: gap of unknown length

* 4868: 4867: gap of 3530 bp in length

* 8398: 4868: gap of unknown length

* 8497: 8398: gap of 4670 bp in length

* 13167: 8497: gap of unknown length

* 13267: 13167: gap of 4523 bp in length

* 17890: 13267: gap of unknown length

* 17891: 17890: gap of 3666 bp in length

* 23556: 17891: gap of unknown length

* 23557: 23556: gap of 9343 bp in length

* 32999: 23557: gap of unknown length

* 33000: 32999: gap of 11468 bp in length

* 44567: 33000: gap of unknown length

* 44568: 44567: gap of 11792 bp in length

* 56459: 44568: gap of unknown length

* 56460: 56459: gap of 15570 bp in length

* 72129: 56460: gap of unknown length

* 72230: 72129: gap of 15644 bp in length

* 87973: 72230: gap of unknown length

* 87974: 87973: gap of 20863 bp in length

* 108936: 87974: gap of unknown length

* 108937: 108936: gap of 21700 bp in length

* 130736: 108937: gap of unknown length

* 151003: 130736: contig of 20267 bp in length

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 177899)

Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Boultard, G.G., Dietrich, N.L., Eagle, N.O., Gupta, J., Ho, S.-L., Huang, M.C., Iod, J., Lee, H.S., Q. Maduro, O.L., Maduro, V.B., Mastrian, S.D., McCloskey, J.C., Ojeda, M.A., Pearson, R., Stantrilop, S., Summers, T.J., Thomas, W., Thomas, P.J., Tongsom, E.E., Touchman, J.W., Tran, J.T., Vogt, J.L., Walker, M.A., Wetherby, K.D. and Green, E.D.

NISC Mouse Sequencing Initiative

Unpublished

2 (bases 1 to 177899)

Green, E.D.

Direct Submission

Submitted (15-JUL-2000) NIH Intramural Sequencing Center, 8717

Grovenmont Circle, Gaithersburg, MD 20877, USA

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.mouse@nih.gov

----- Project Information

Center project name: vc

Center clone name: 348B11

----- Summary statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 171084 bases at least Q40

Consensus quality: 172951 bases at least Q30

Consensus quality: 174122 bases at least Q20

Insert size: 187000; agarose-fp

Insert size: 202000; pulse-field-gel

Insert size: 176499; sum-of-contigs

Quality coverage: 4.87x in Q20 bases; agarose-fp

Quality coverage: 4.50x in Q20 bases; pulse-field-gel

Quality coverage: 5.16x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 2463: contig of 2463 bp in length

* 2464: 2463: gap of unknown length

* 2465: 2464: contig of 2204 bp in length

* 4867: 2465: gap of unknown length

* 4868: 4867: gap of 3530 bp in length

* 8398: 4868: gap of unknown length

* 8497: 8398: gap of 4670 bp in length

* 13167: 8497: gap of unknown length

* 13267: 13167: gap of 4523 bp in length

* 17890: 13267: gap of unknown length

* 17891: 17890: gap of 3666 bp in length

* 23556: 17891: gap of unknown length

* 23557: 23556: gap of 9343 bp in length

* 32999: 23557: gap of unknown length

* 33000: 32999: gap of 11468 bp in length

* 44567: 33000: gap of unknown length

* 44568: 44567: gap of 11792 bp in length

* 56459: 44568: gap of unknown length

* 56460: 56459: gap of 15570 bp in length

* 72129: 56460: gap of unknown length

* 72230: 72129: gap of 15644 bp in length

* 87973: 72230: gap of unknown length

* 87974: 87973: gap of 20863 bp in length

* 108936: 87974: gap of unknown length

* 108937: 108936: gap of 21700 bp in length

* 130736: 108937: gap of unknown length

* 151003: 130736: contig of 20267 bp in length

Fri Dec 28 08:22:45 2001

us-09-830-647-3.rge

Page 24

